Machine Learning for Design

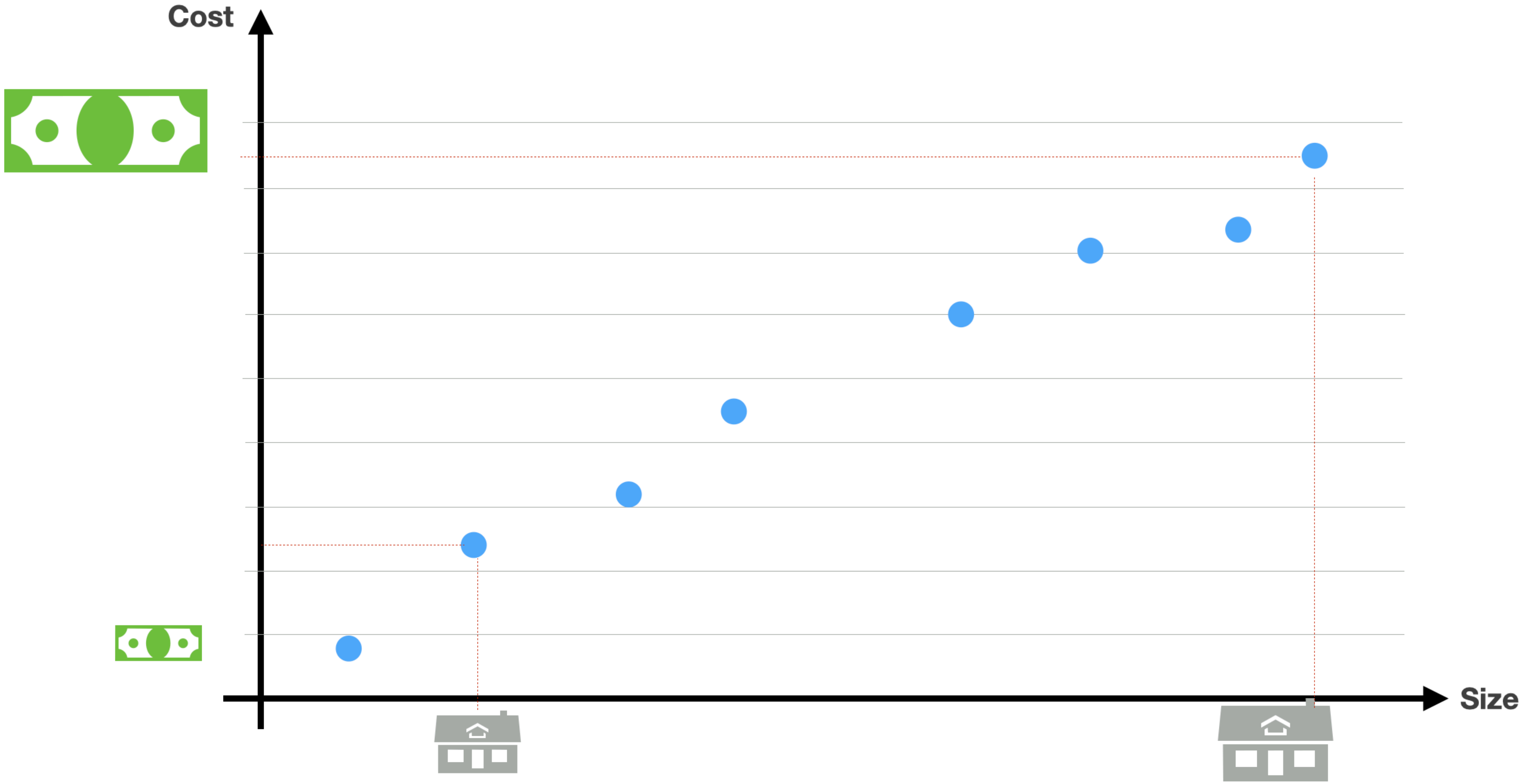
Lecture 3

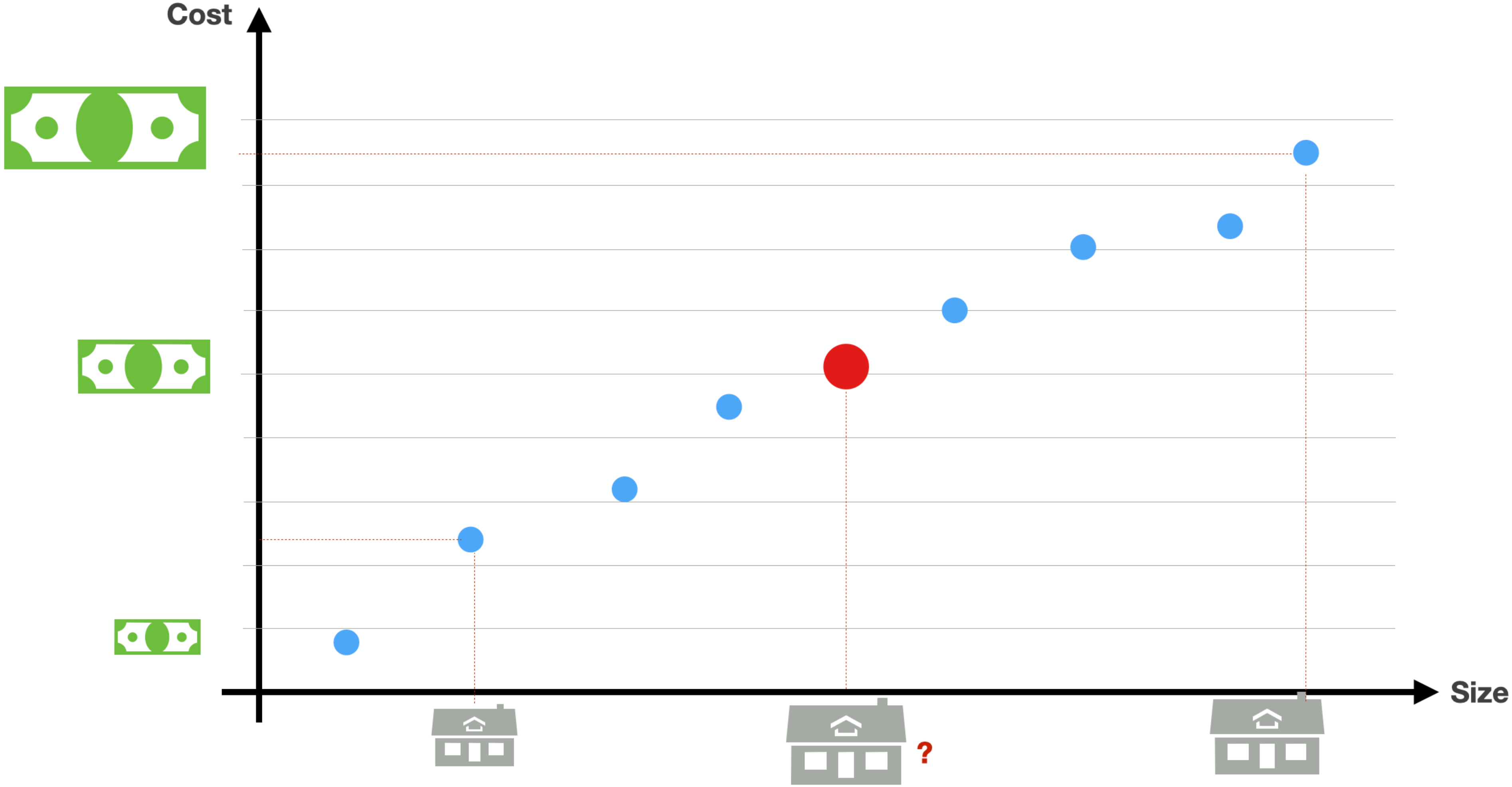
Machine Learning for Images. *Part 1*

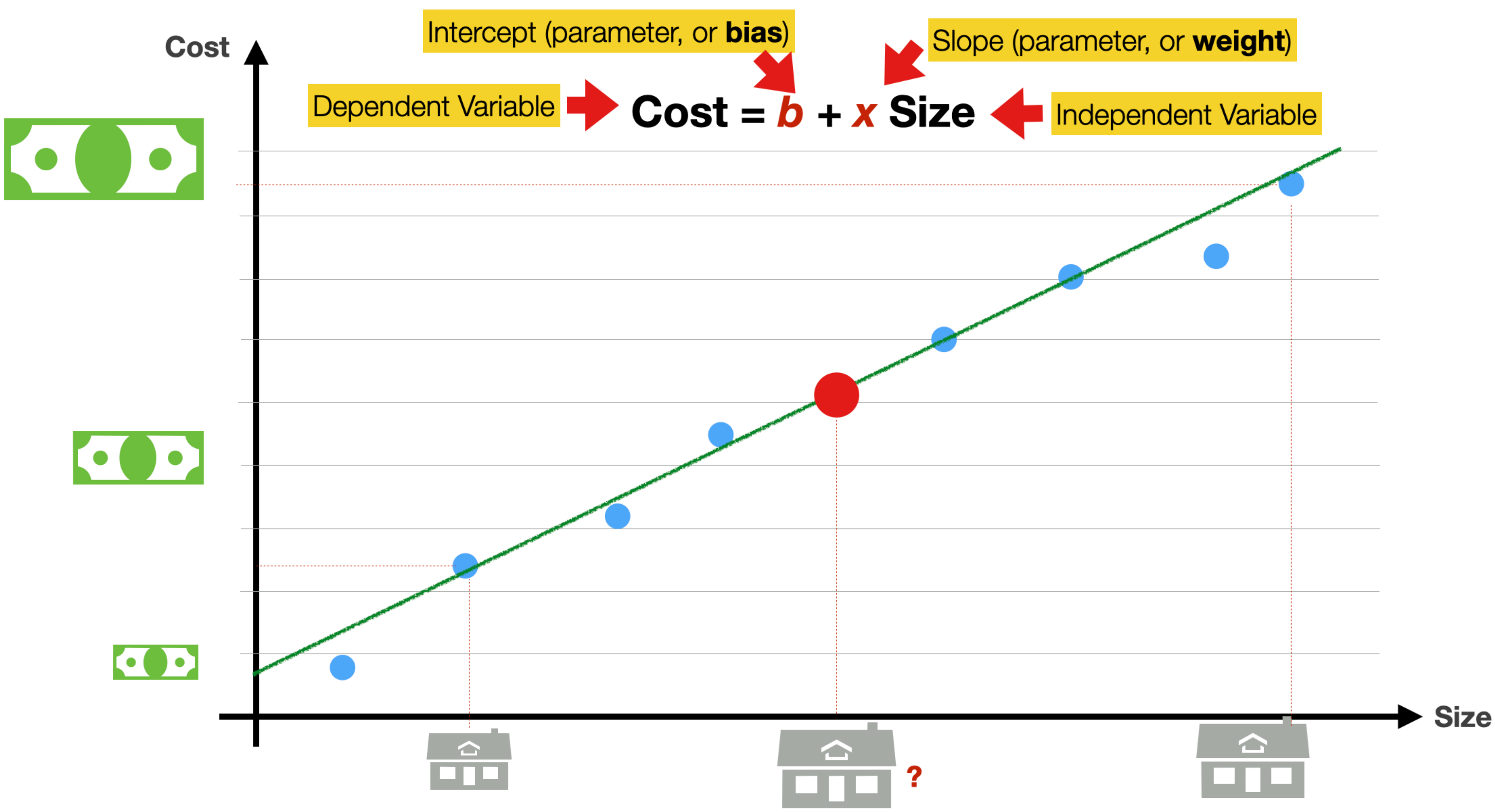
A bit more on regression and classification

And your very first contact with (deep) neural networks

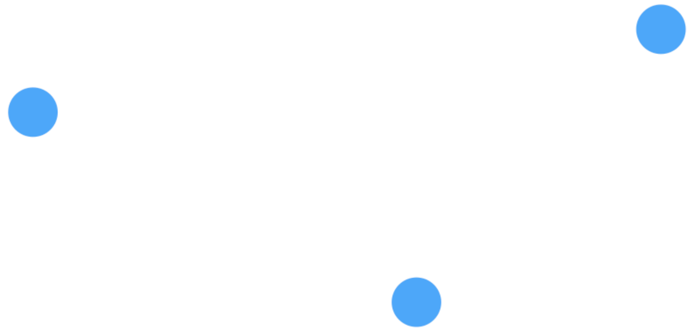
Linear Regression


Cost
Size

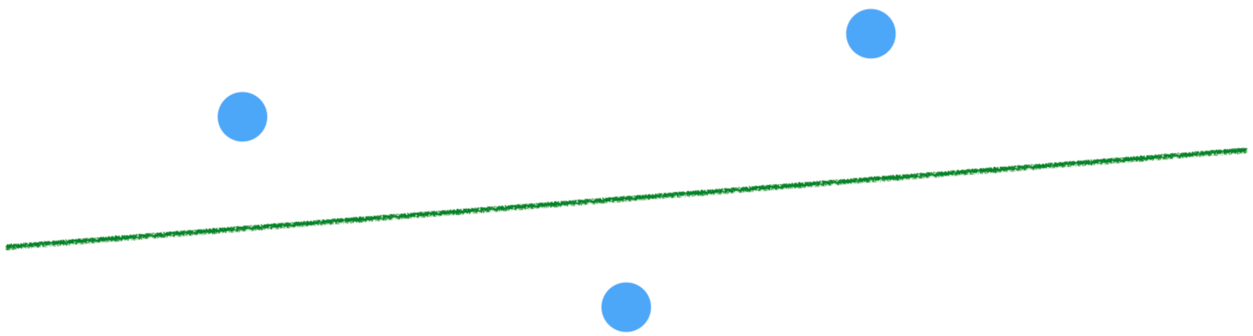

Cost
Size
?

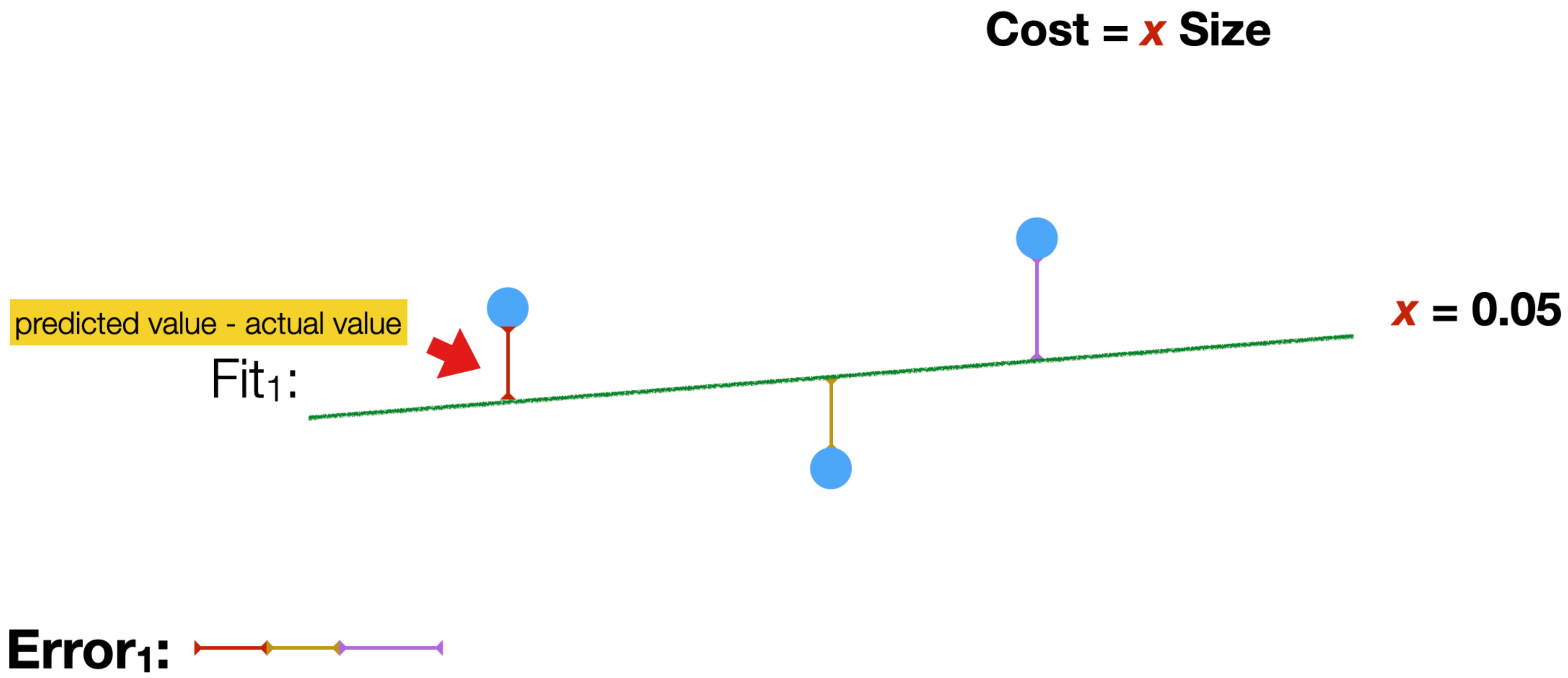

Intercept (parameter, or bias)
Slope (parameter, or weight)
Cost
Dependent Variable
Cost = b + x Size
Independent Variable
Size
?

Cost = *x* Size



Cost = *x* Size

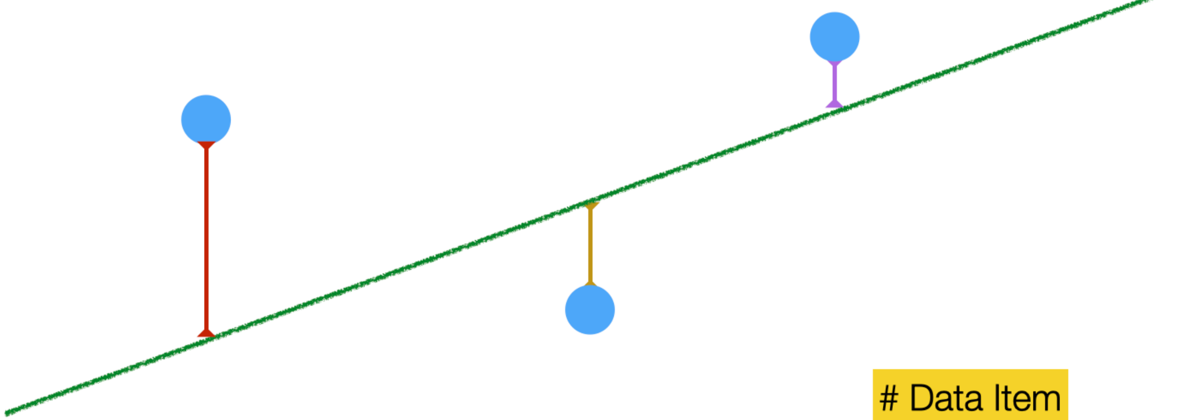



Cost = x Size
x = 0.05
predicted value - actual value
Fit1:
Error1:

Cost = *x* Size

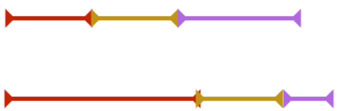
***x* = 0.2**

Fit$_2$:



Error$_1$:

Error$_2$:



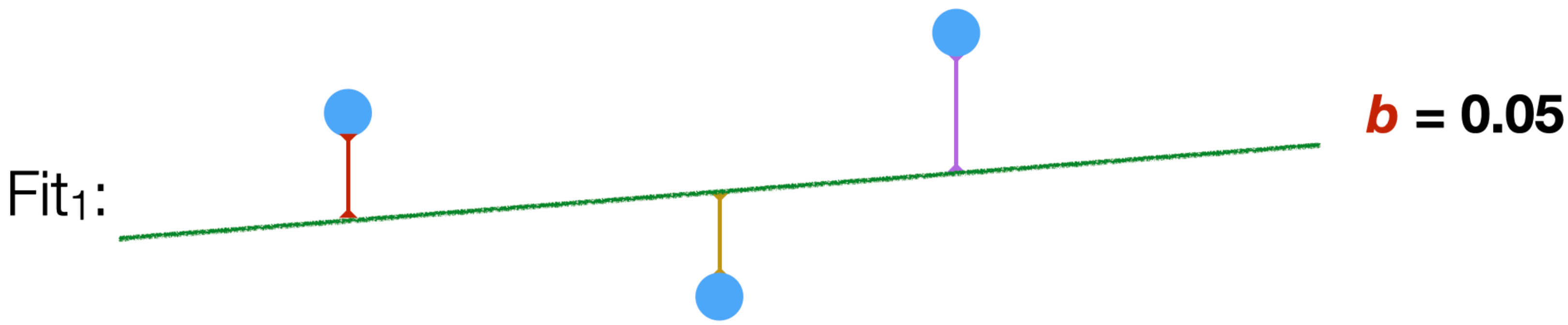
\# Data Item

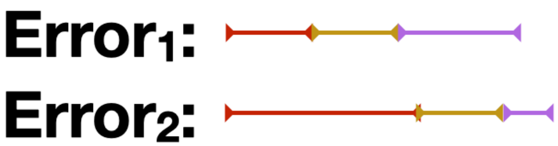
$$Error = \frac{1}{2d}\sum_{i=1}^{d}(Prediction_i - Value_i)^2$$

Mean Squared Error (MSE)

Current data item

Cost = *x* Size





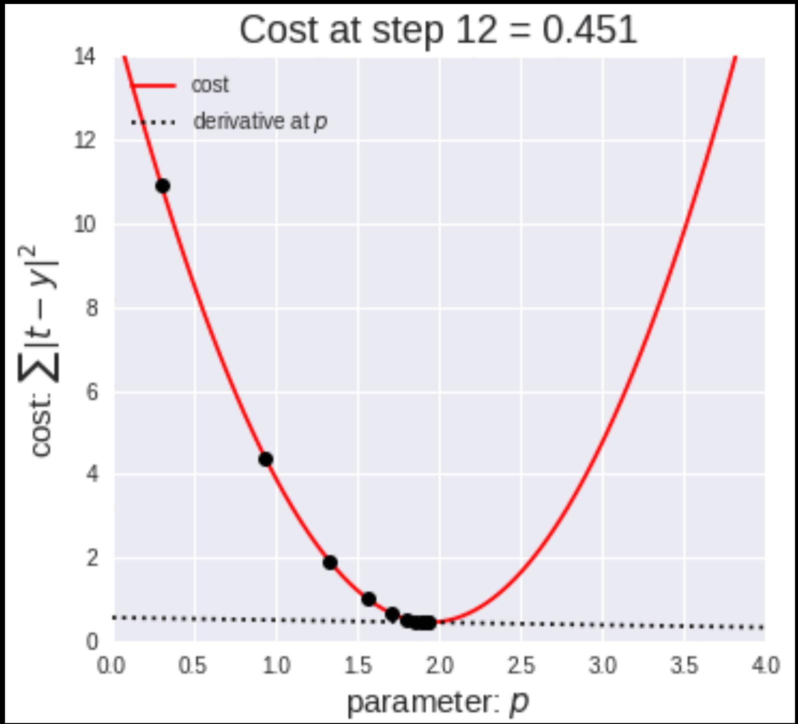
Fit$_1$ is a better fit on the training data than Fit$_2$

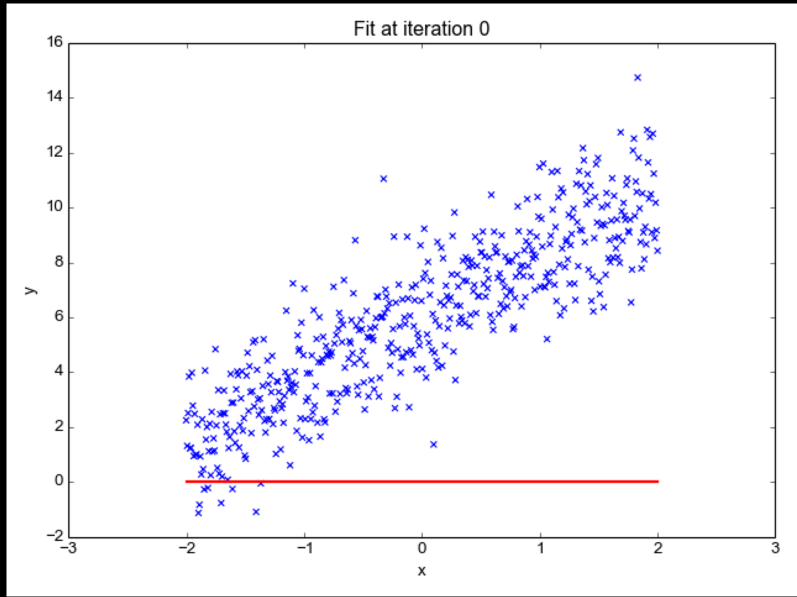
We select ***b*** **= 0.05**

Finding the best parameter values

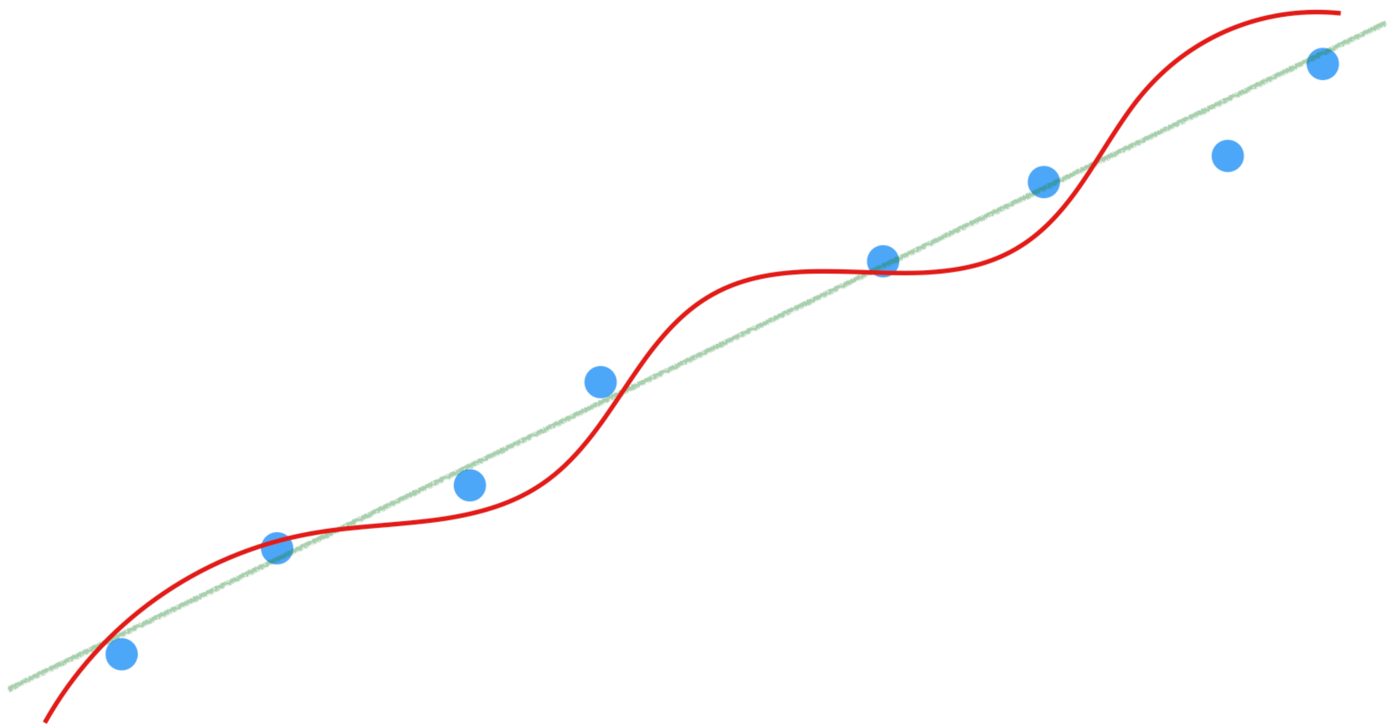
- Training the model

- **Gradient Descent**: an algorithm to find the minimum point of a function
- **Hyperparameters**: parameters of the Gradient Descent
  - *Learning Rate*: speed of descent
  - *Epochs*: max number of steps

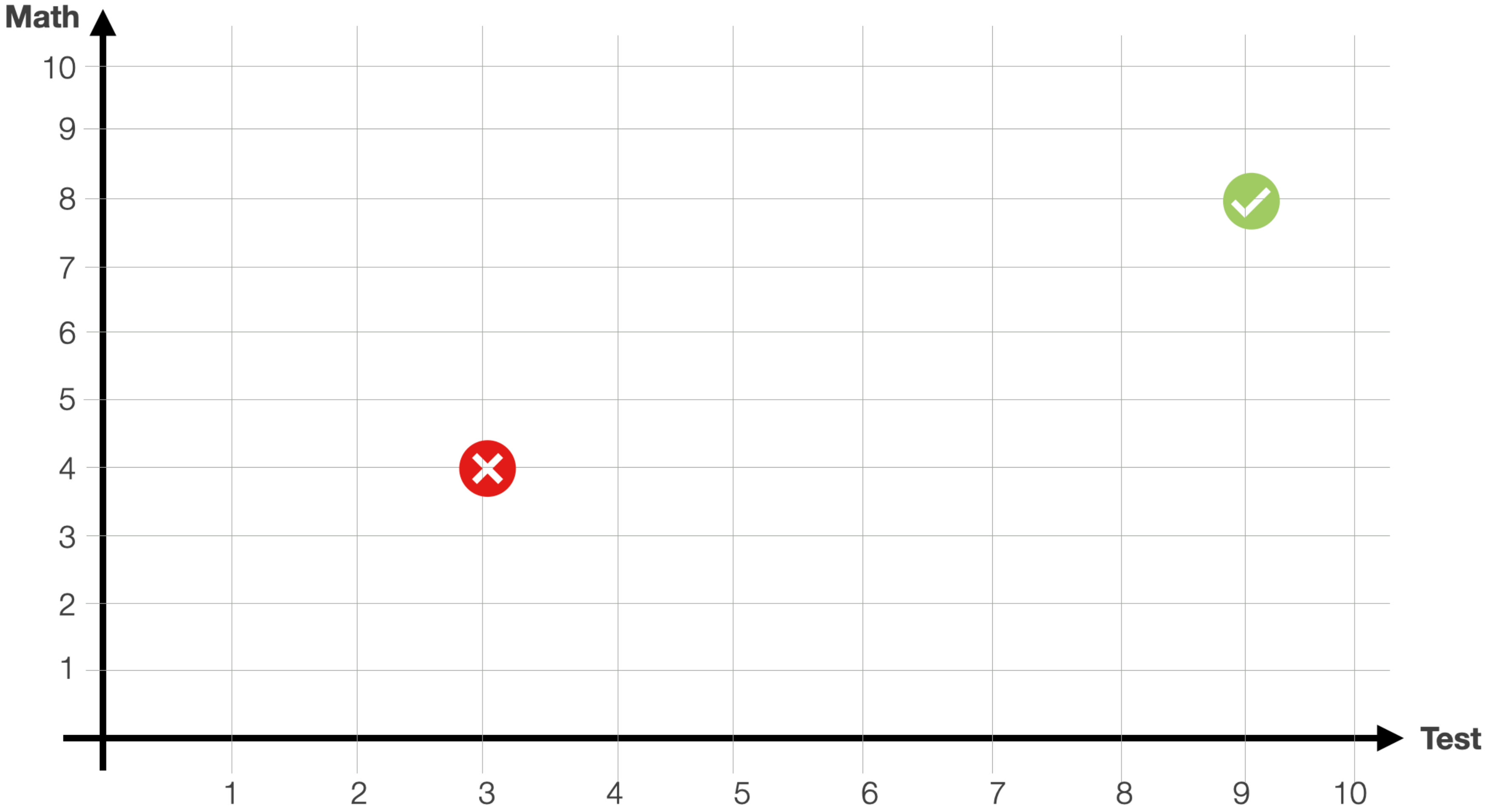

Cost at step 12 = 0.451
cost
derivative at p
cost: ∑|t − y|²
parameter: p
0.0
0.5
1.0
1.5
2.0
2.5
3.0
3.5
4.0
0
2
4
6
8
10
12
14

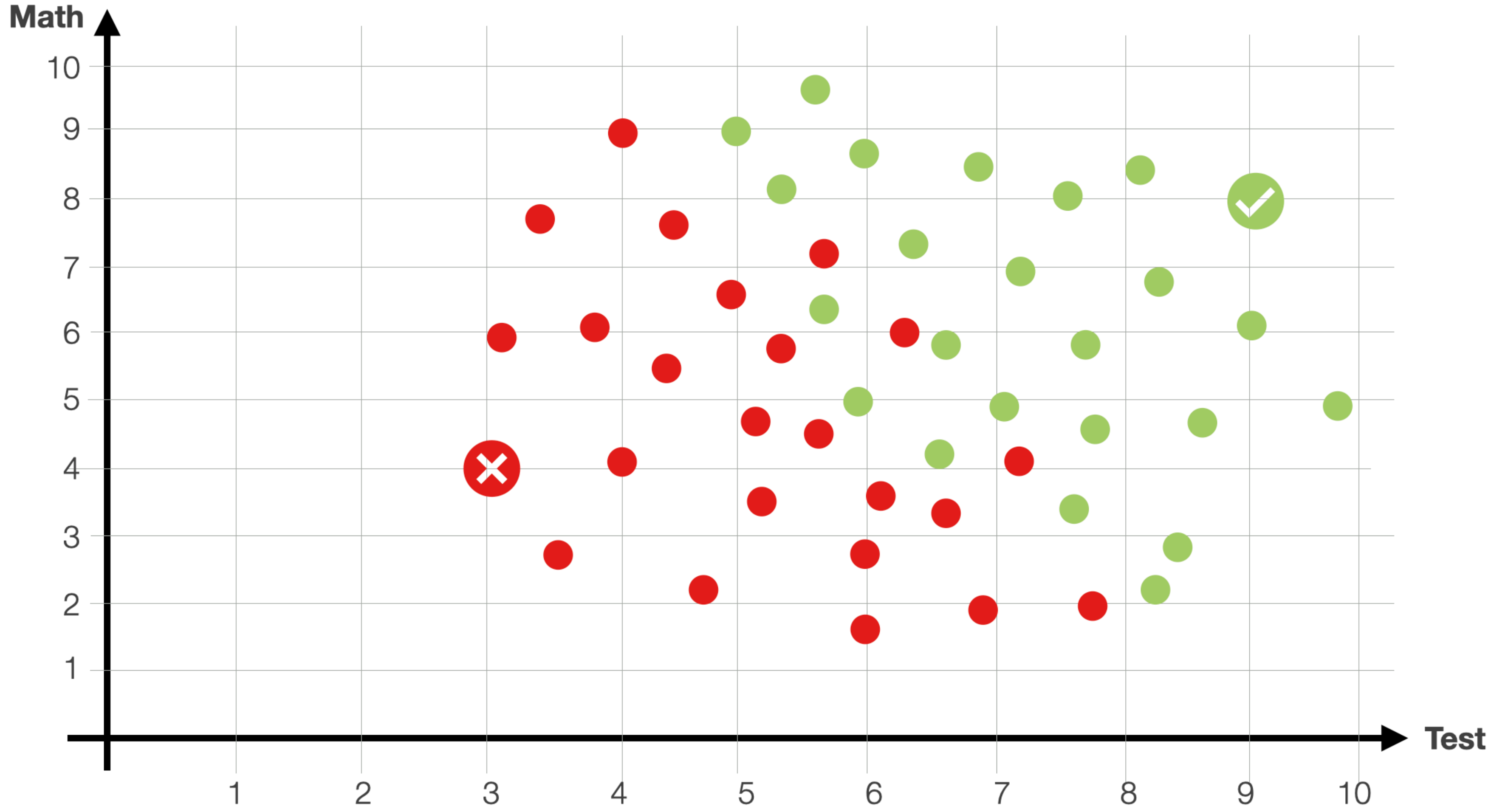

Fit at iteration 0
y
x
16
14
12
10
8
6
4
2
0
−2
−3
−2
−1
0
1
2
3

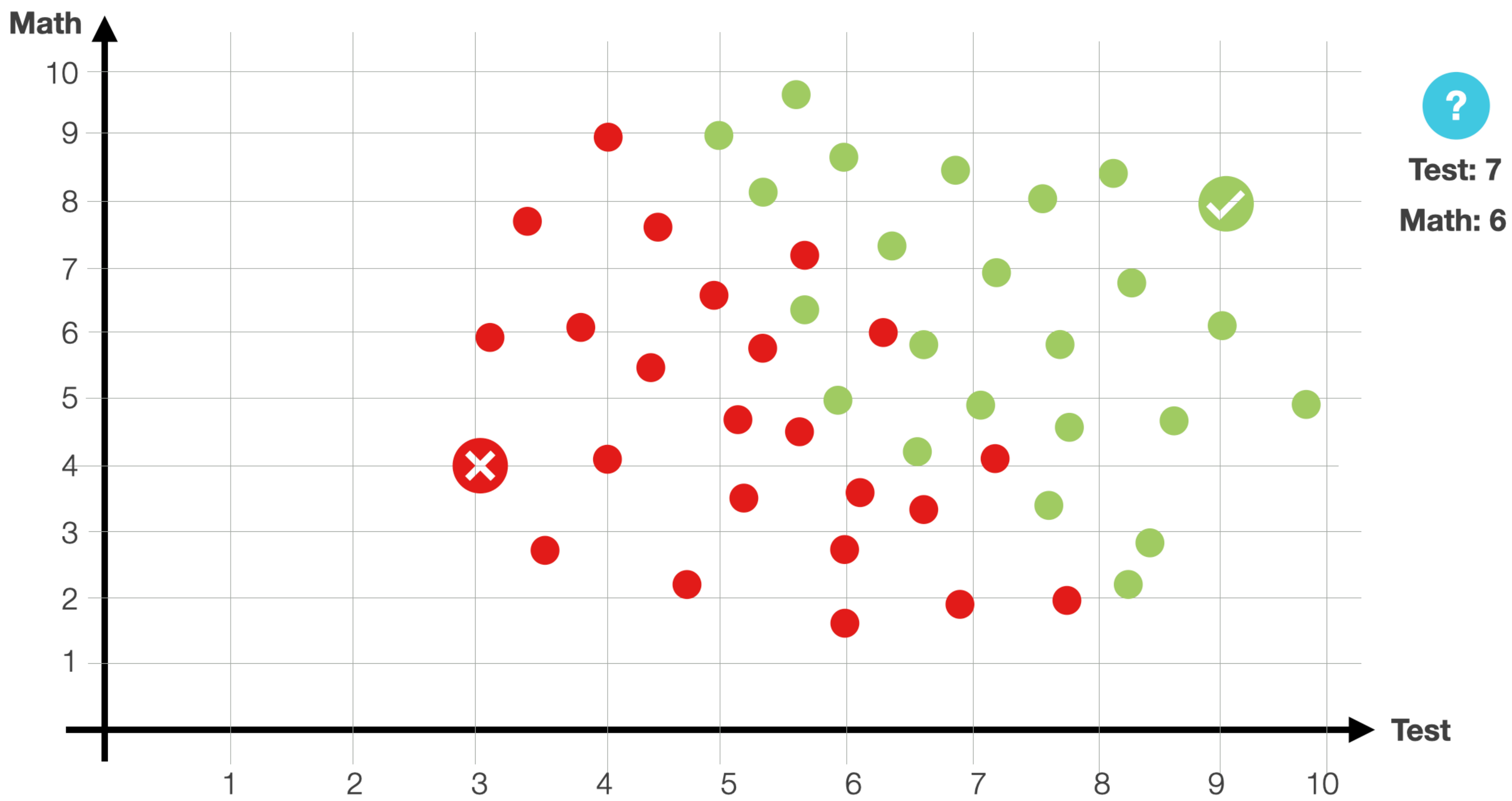
N^{th} degree polynomial → **Cost = b + x_1 Size + x_2 Size^2 + …. + x_n Size^n**



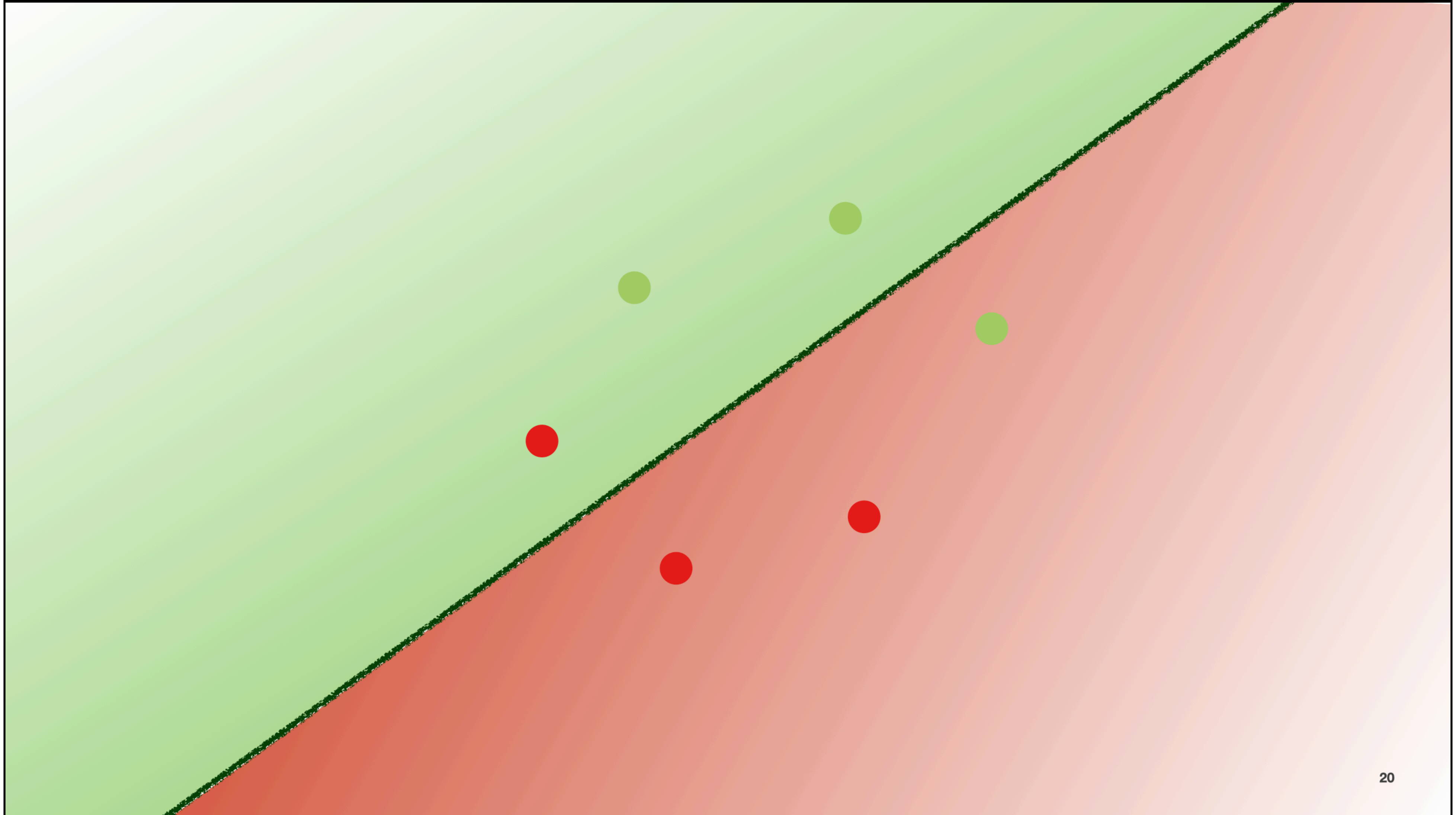
Classification

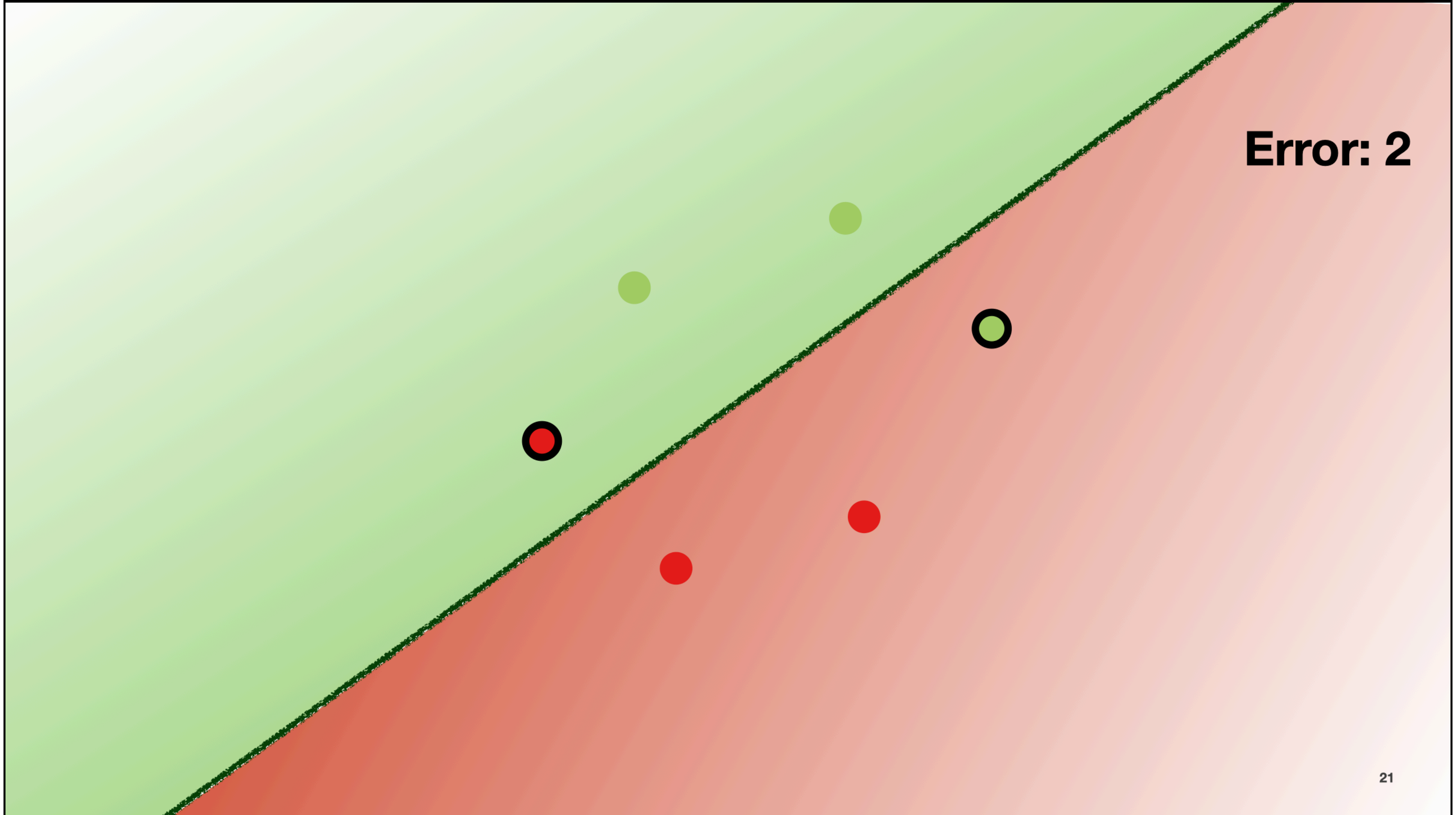

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

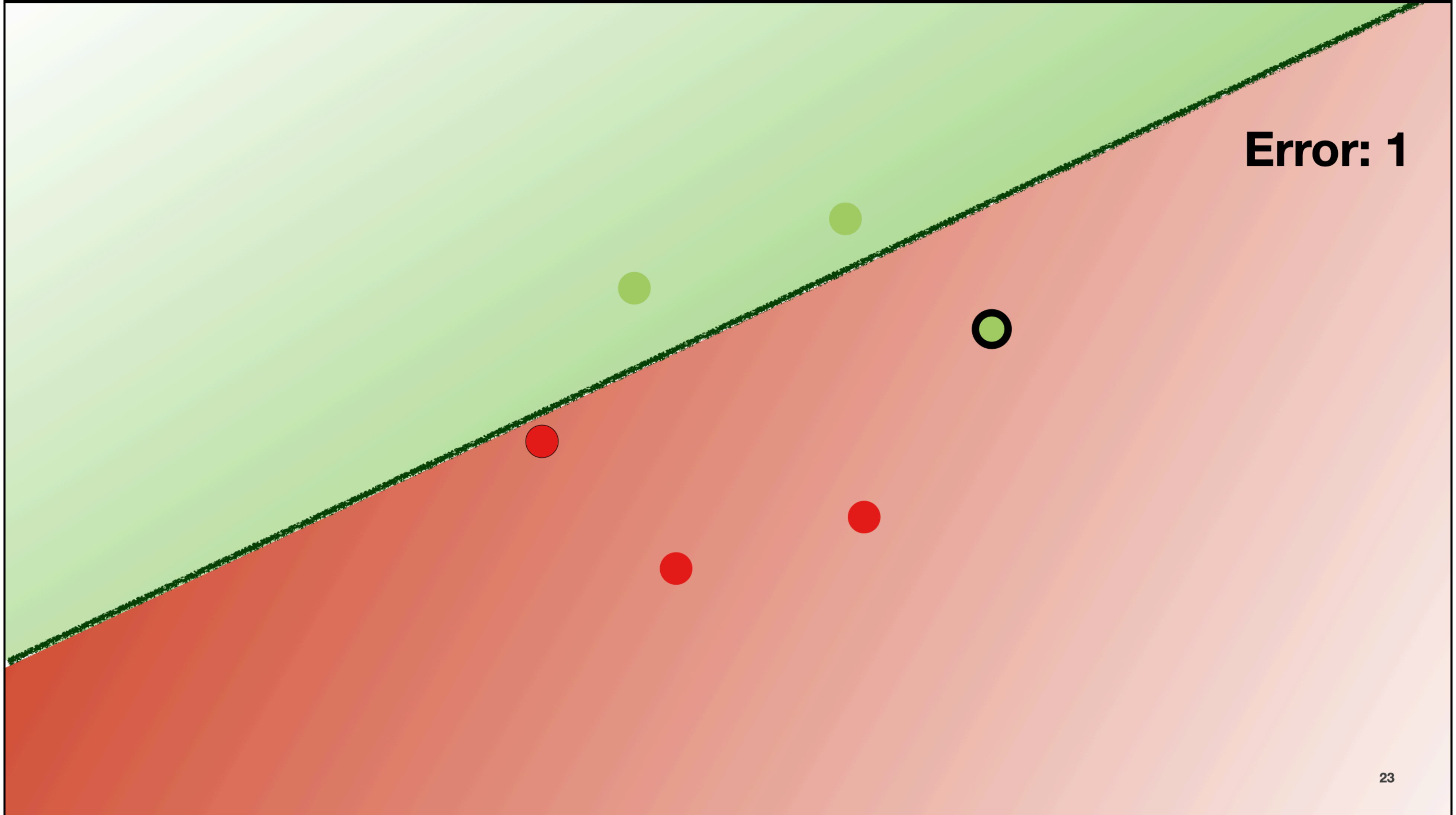

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

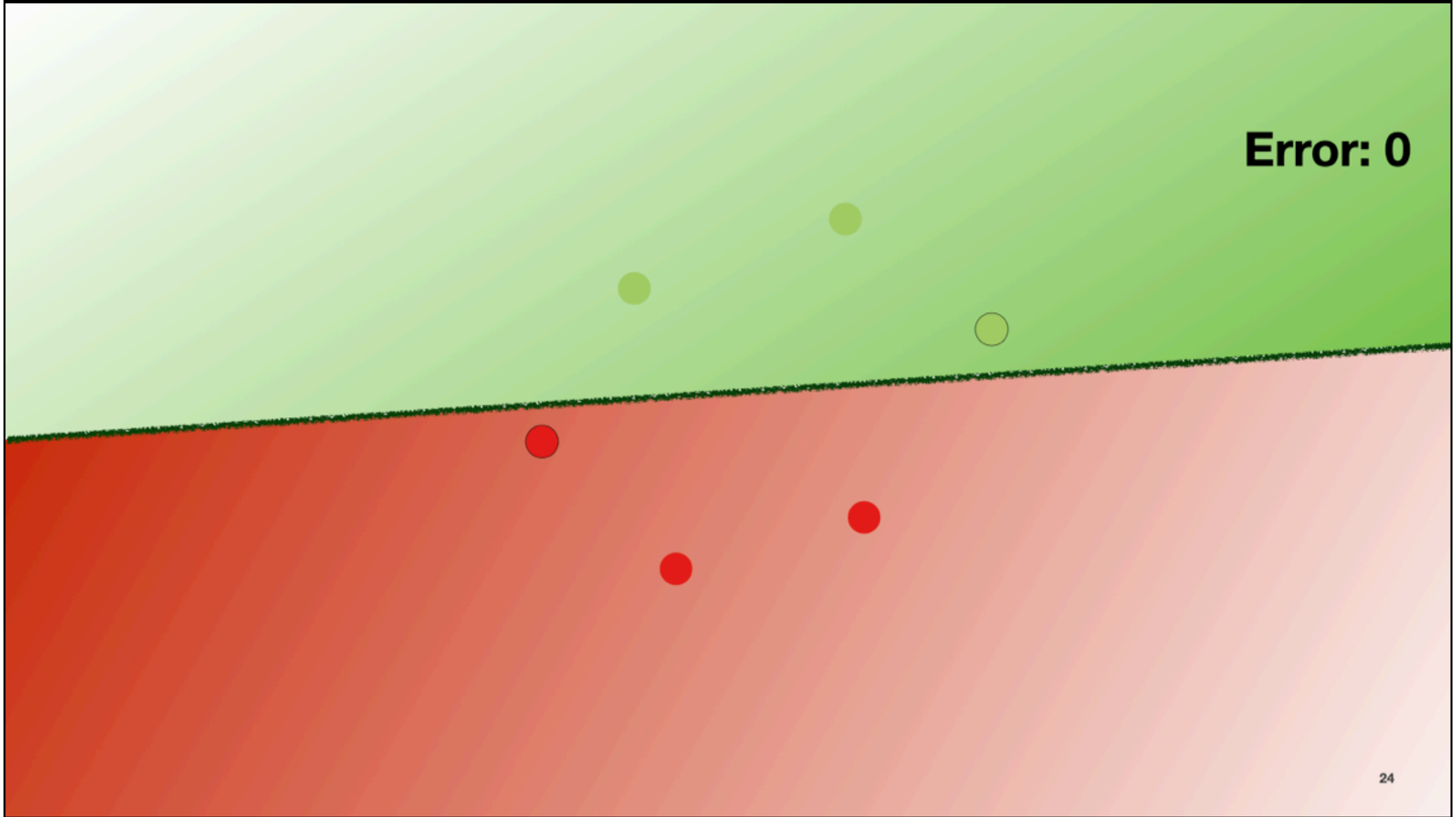

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10
?
Test: 7
Math: 6

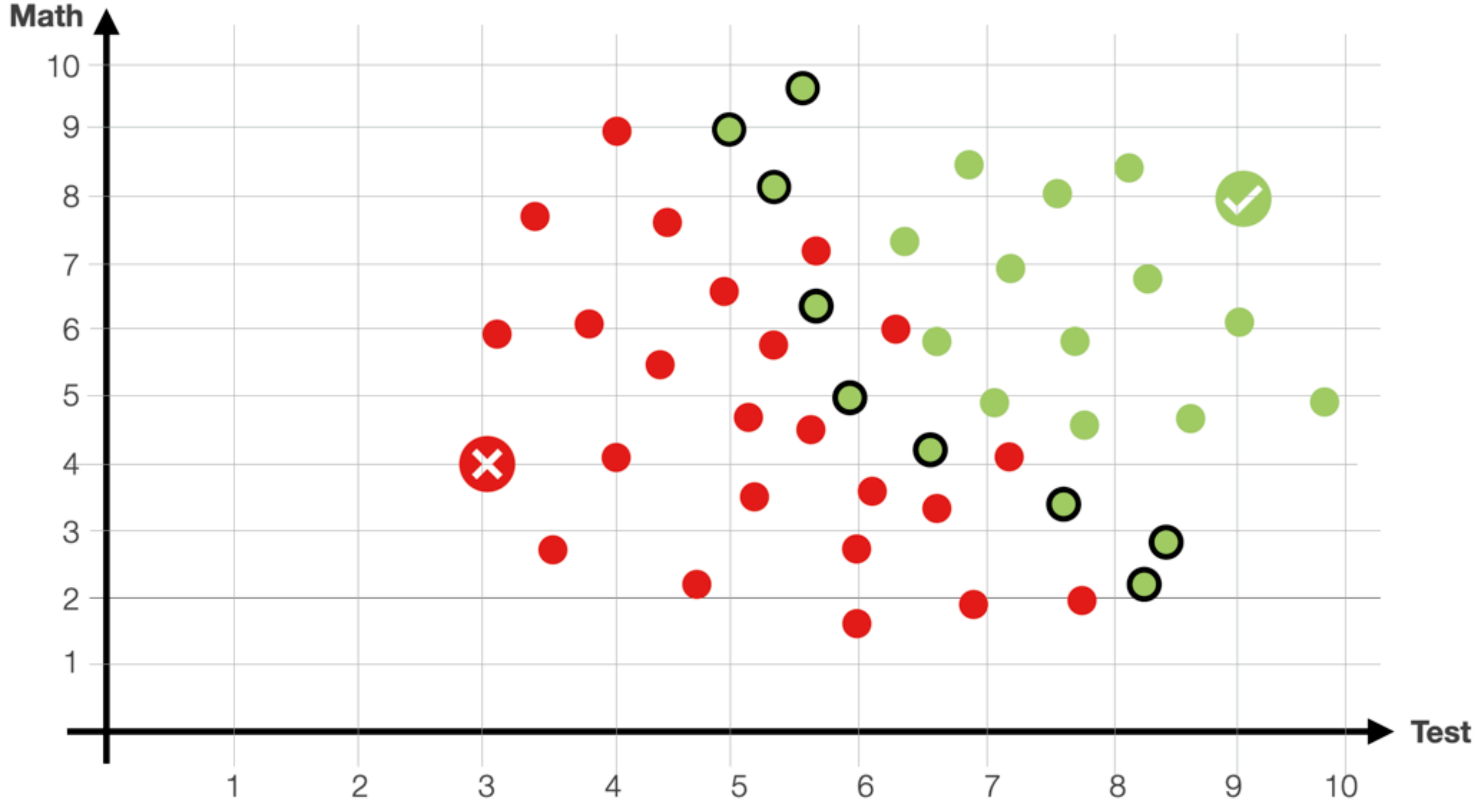
Logistic Regression



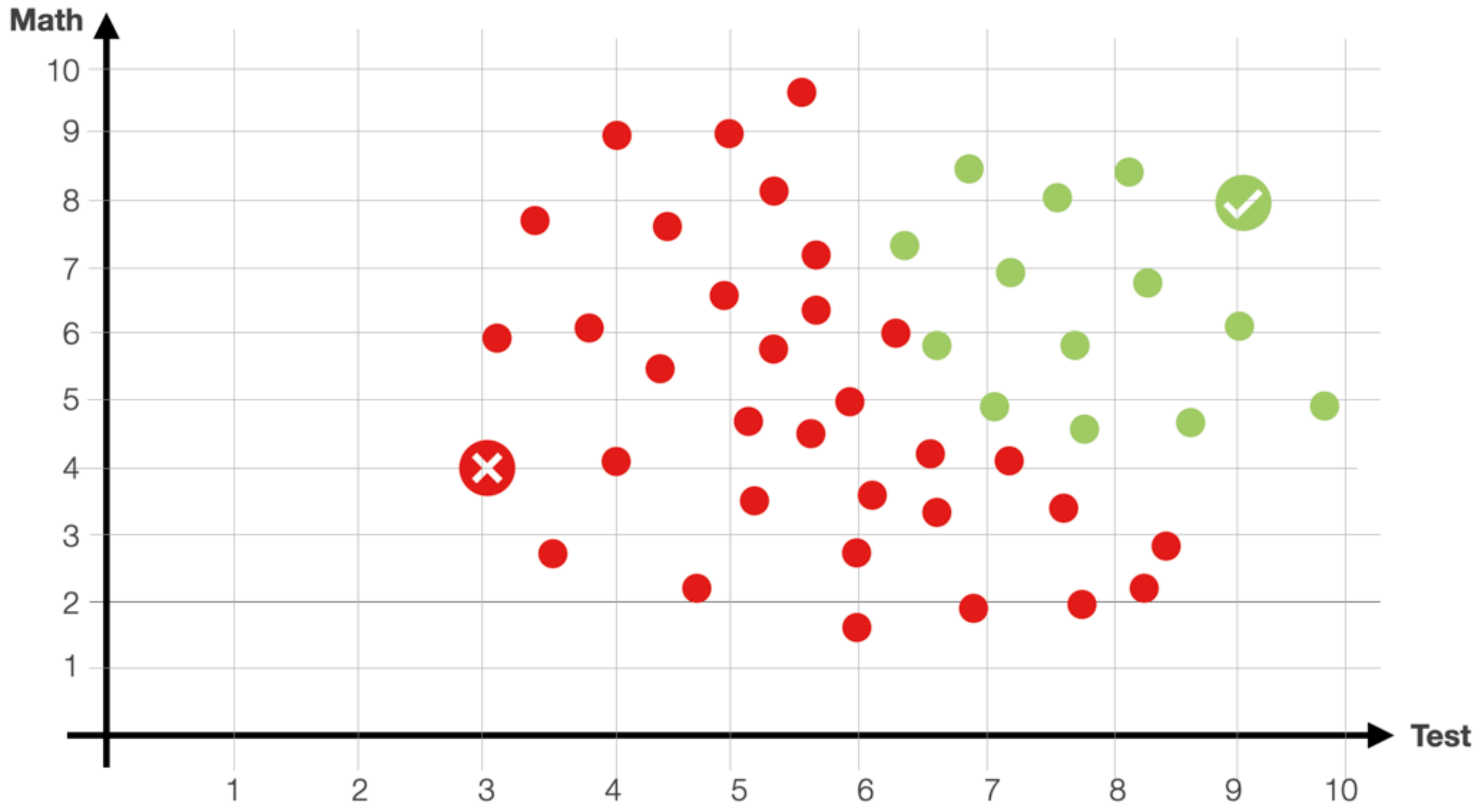

Error: 2

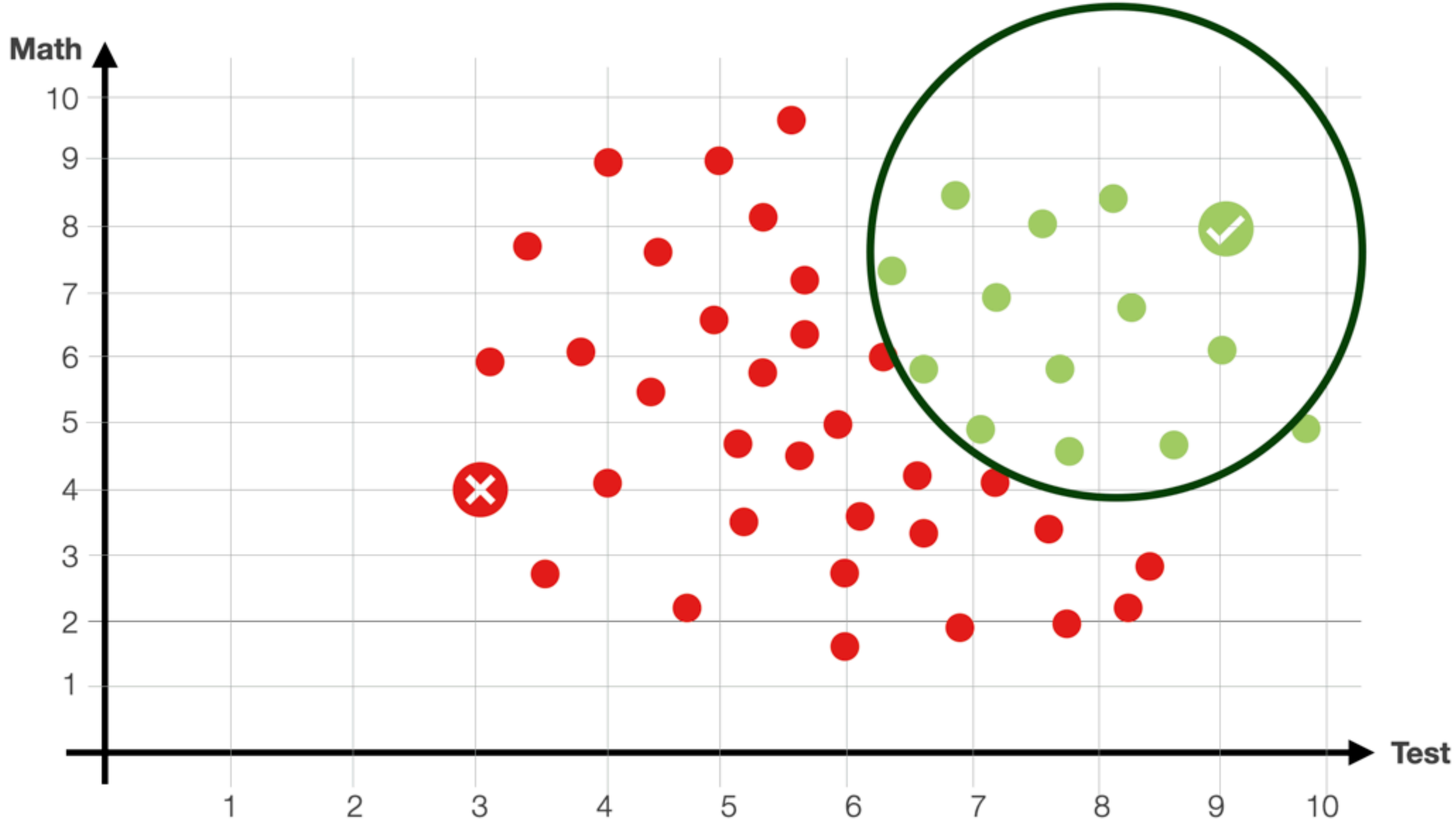

Error: 1
23


Error: 0
24

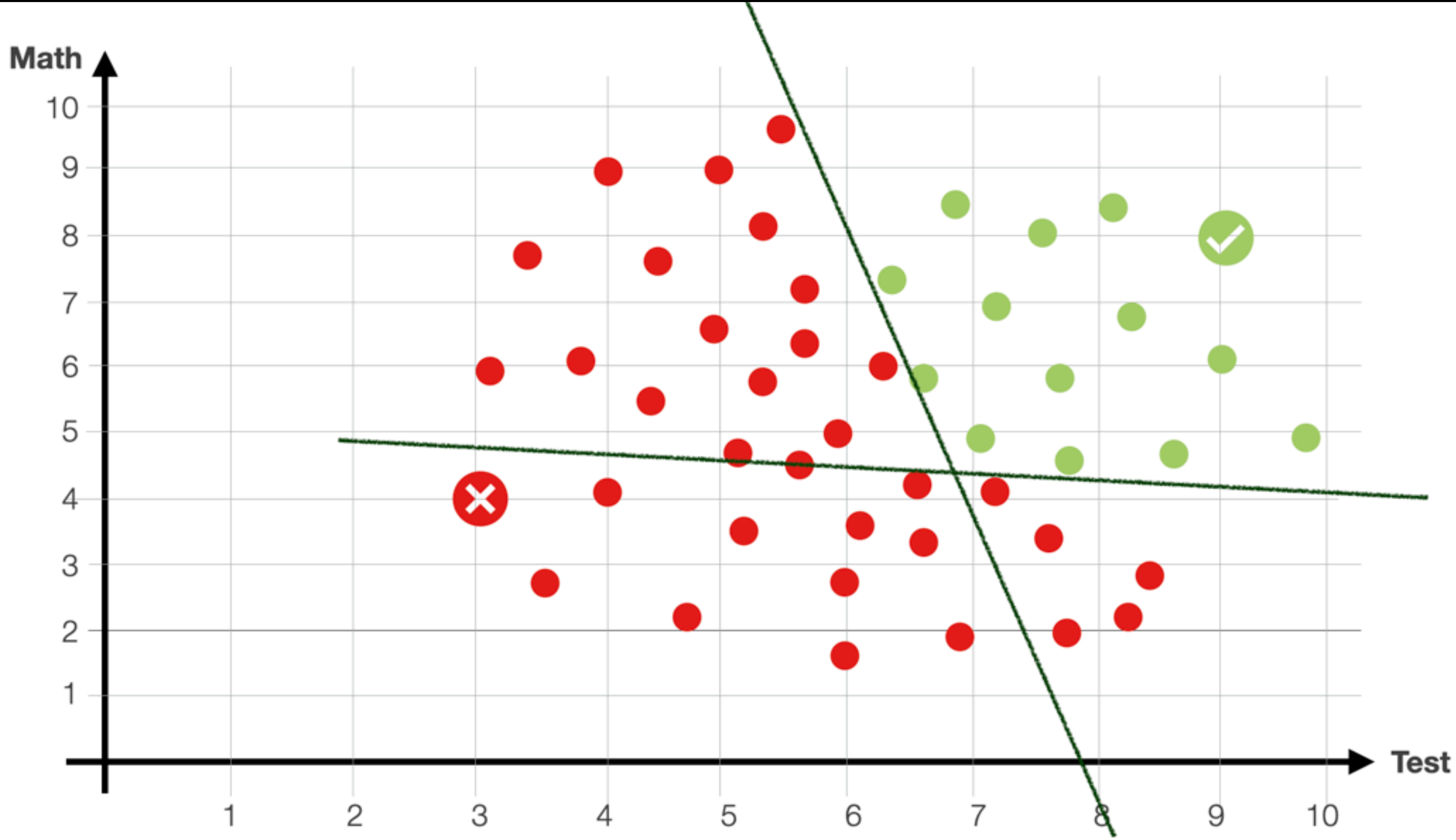


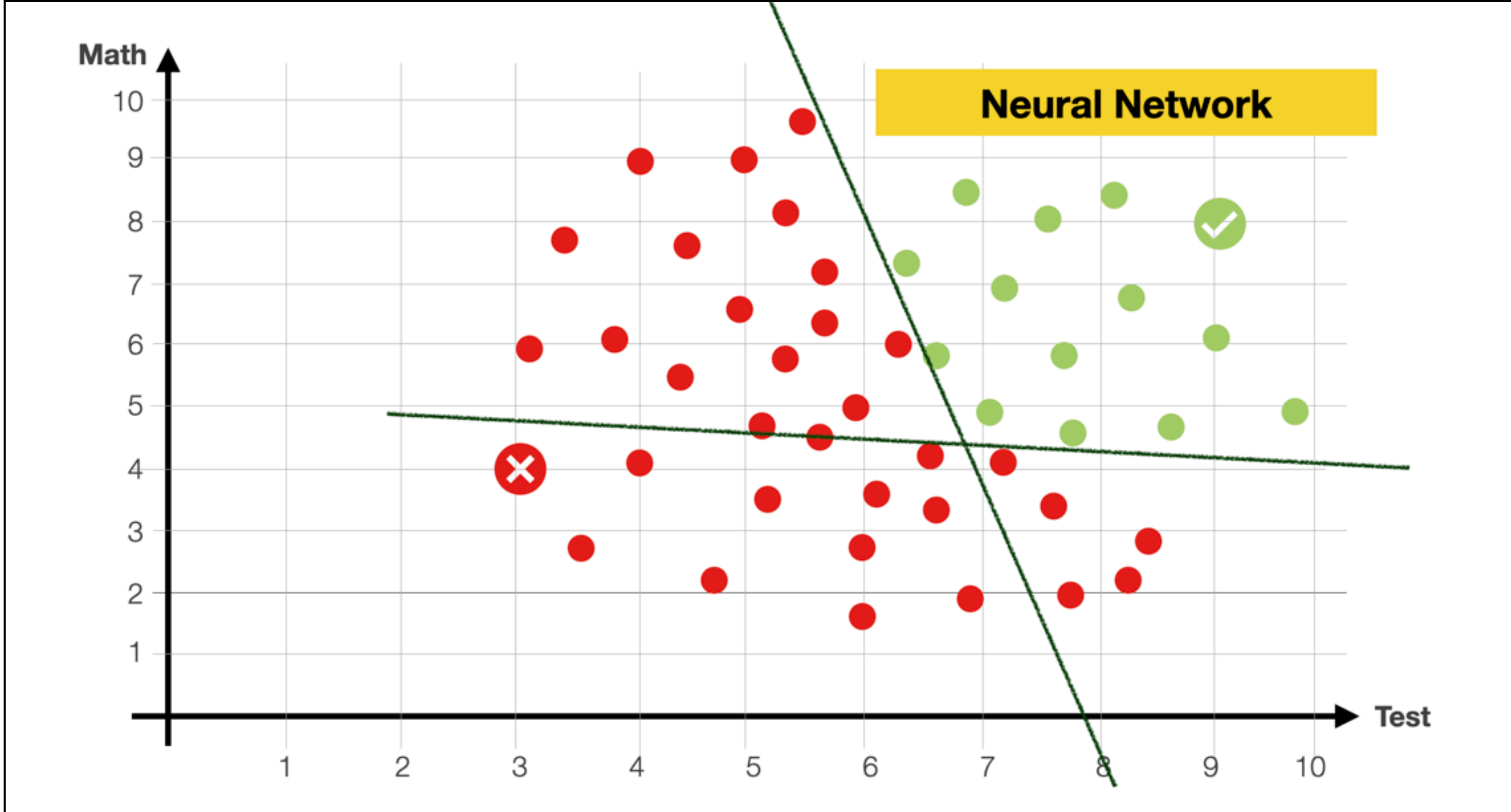
Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

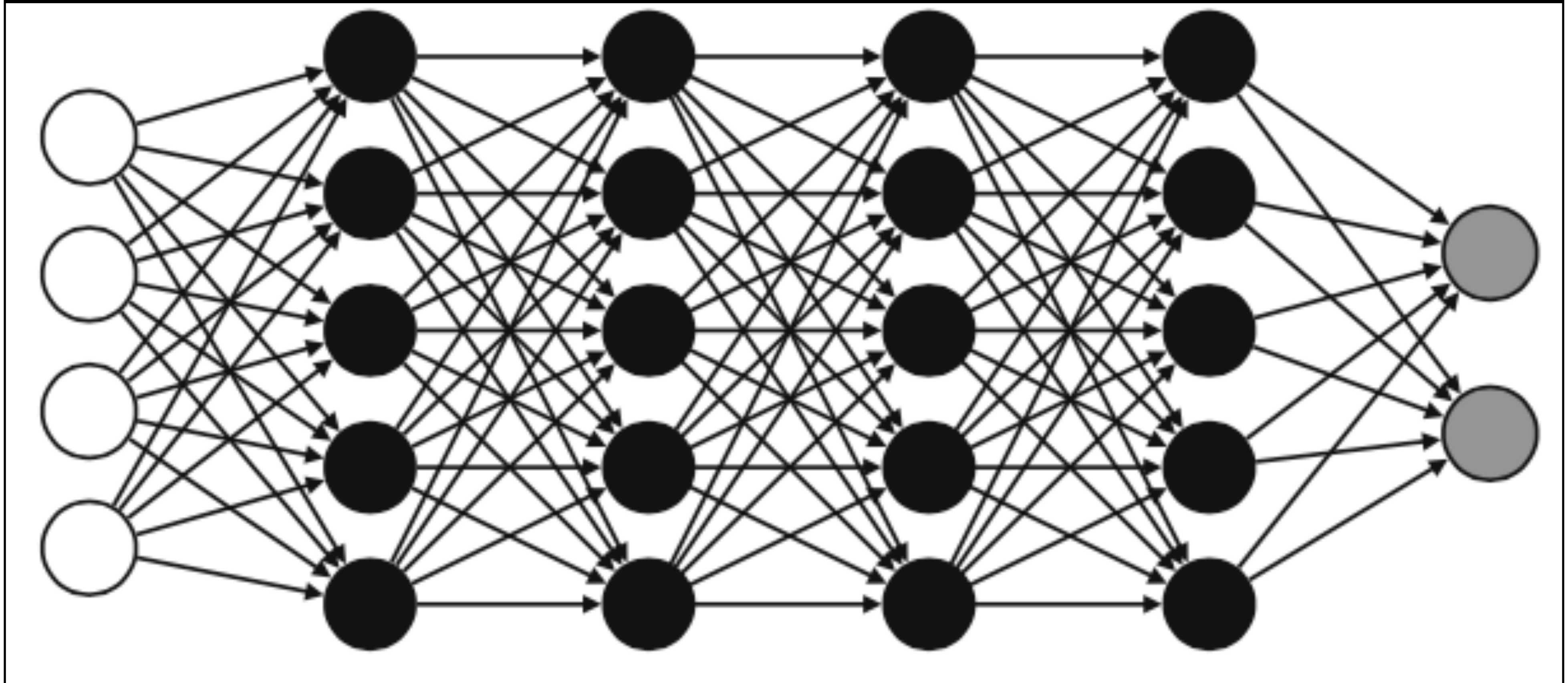

Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10

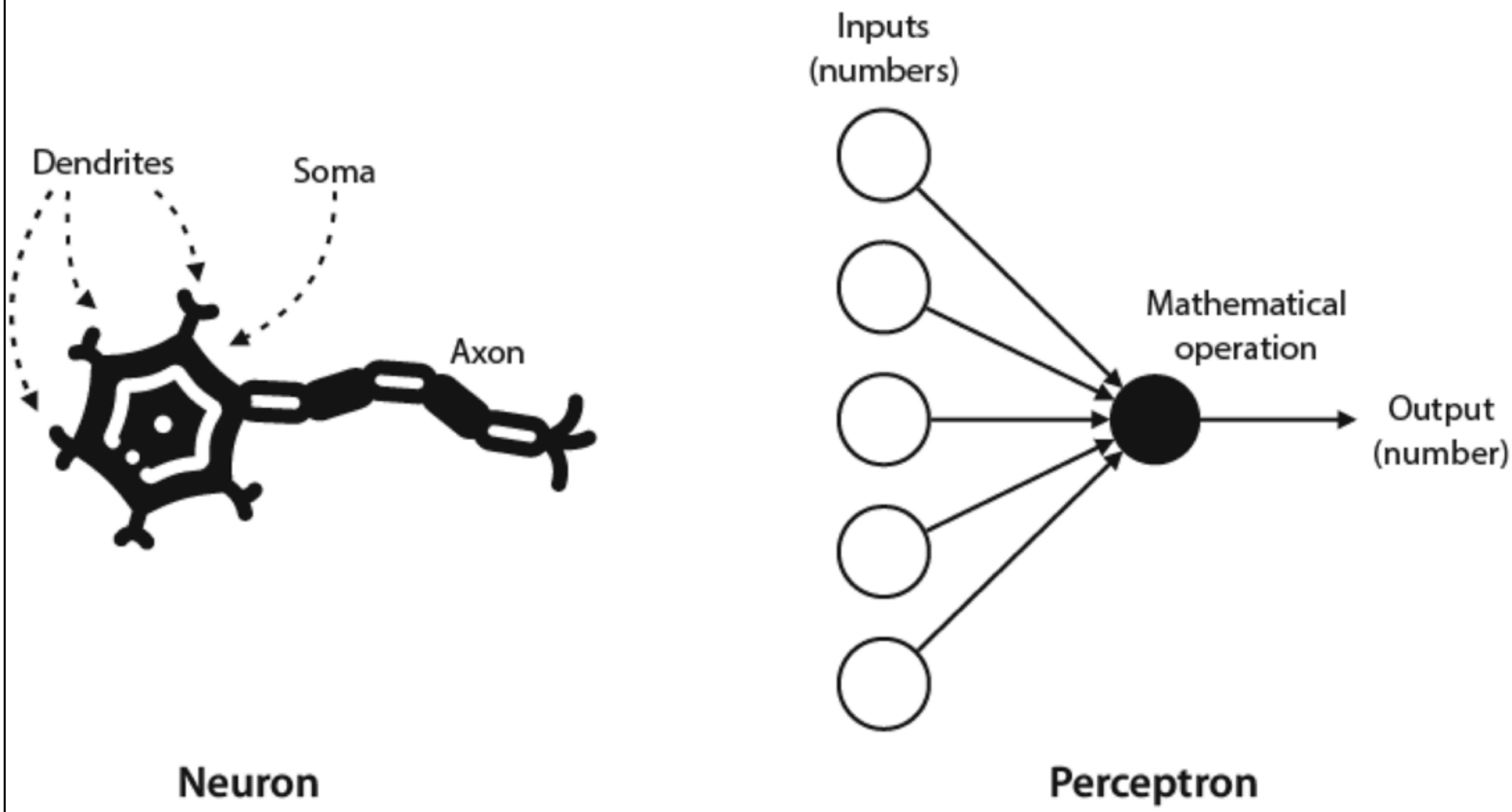


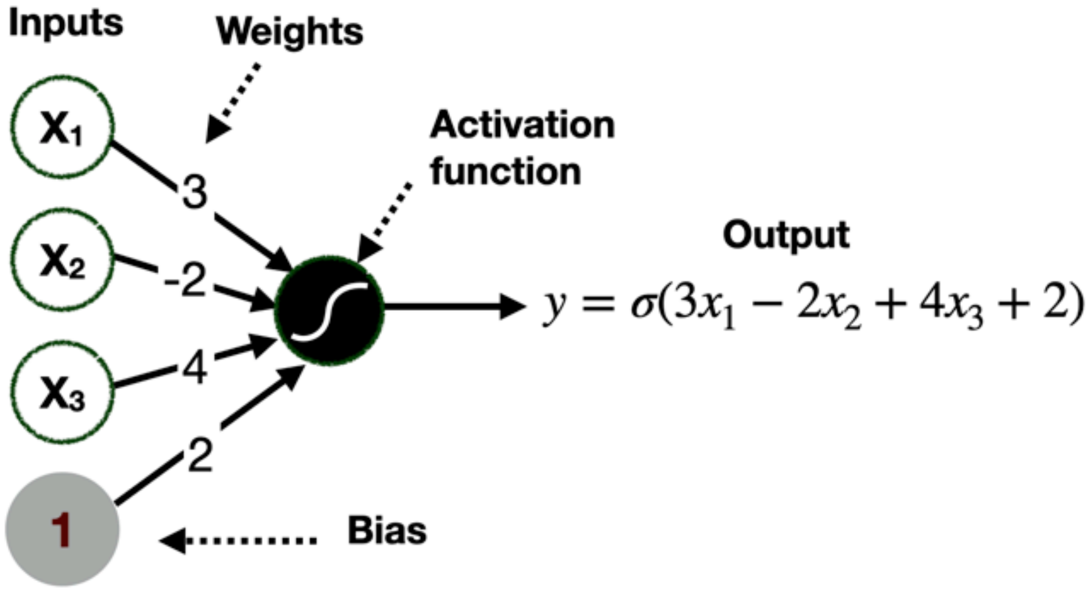
Math
10
9
8
7
6
5
4
3
2
1
Test
1
2
3
4
5
6
7
8
9
10


Math
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10
Test


Neural Network
Math
10
9
8
7
6
5
4
3
2
1
1
2
3
4
5
6
7
8
9
10
Test

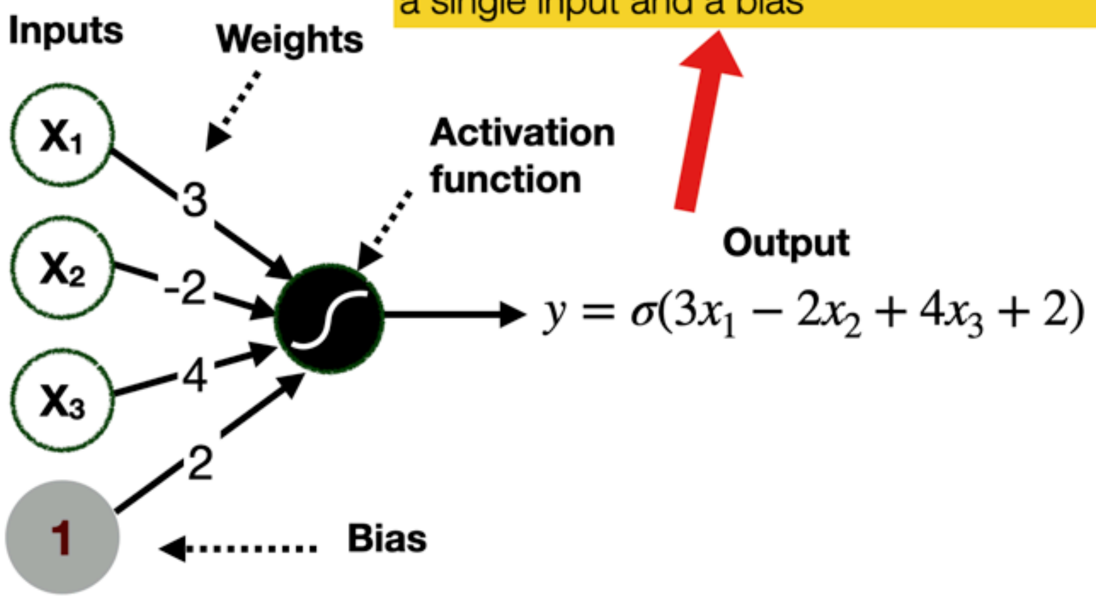


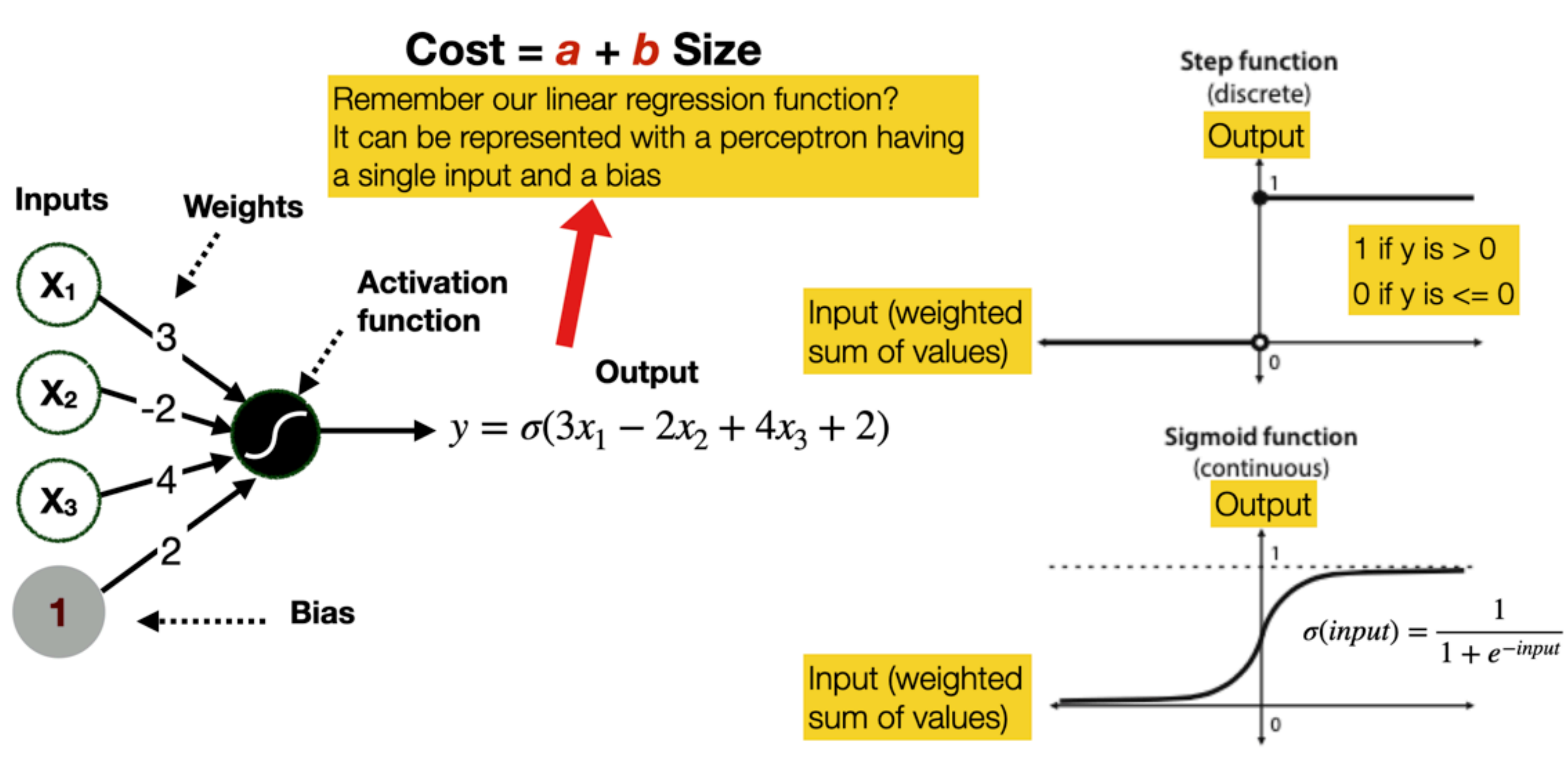

Dendrites
Soma
Axon
Neuron
Inputs
(numbers)
Mathematical
operation
Output
(number)
Perceptron

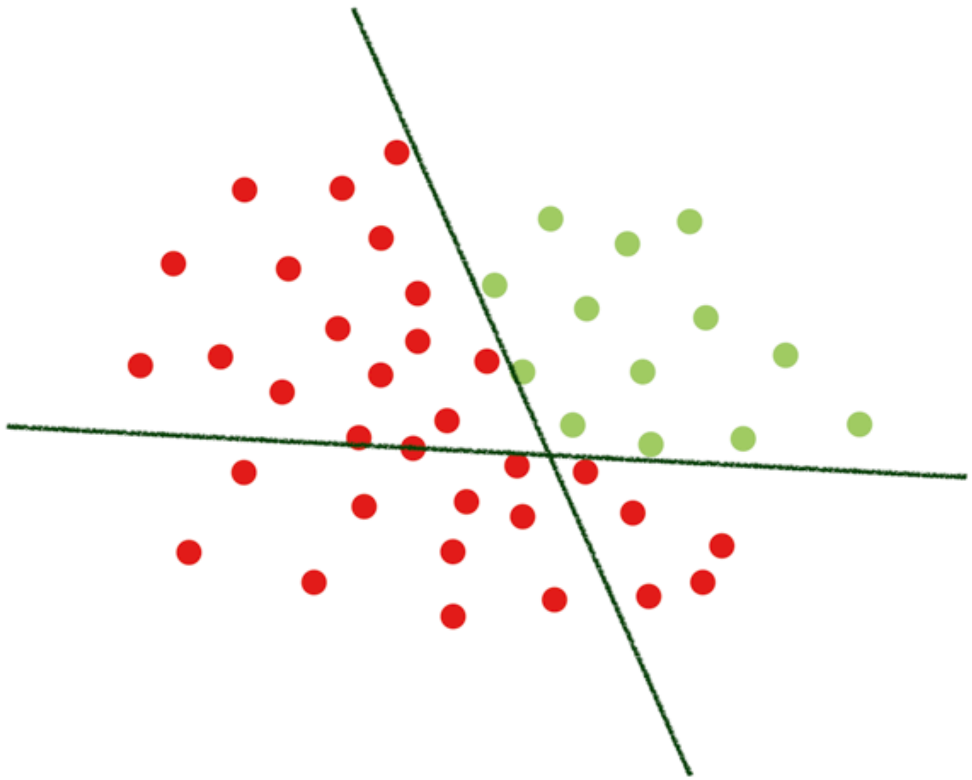

Inputs
Weights
Activation function
Output
X1
X2
X3
1
3
-2
4
2
Bias
$y = \sigma(3x_1 - 2x_2 + 4x_3 + 2)$

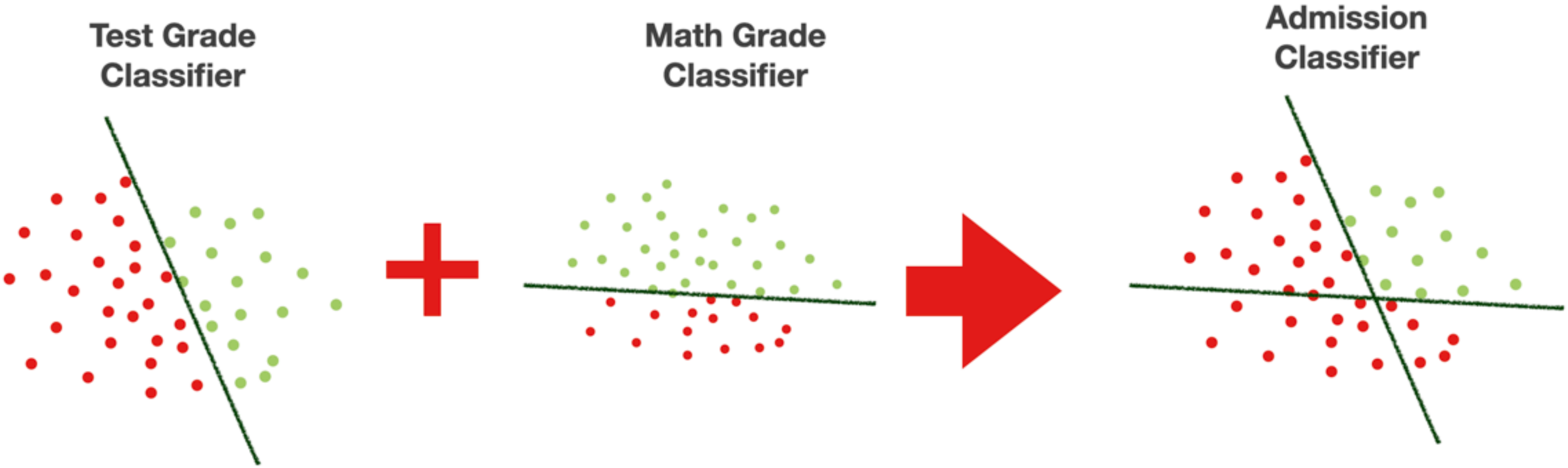
Cost = *a* + *b* Size

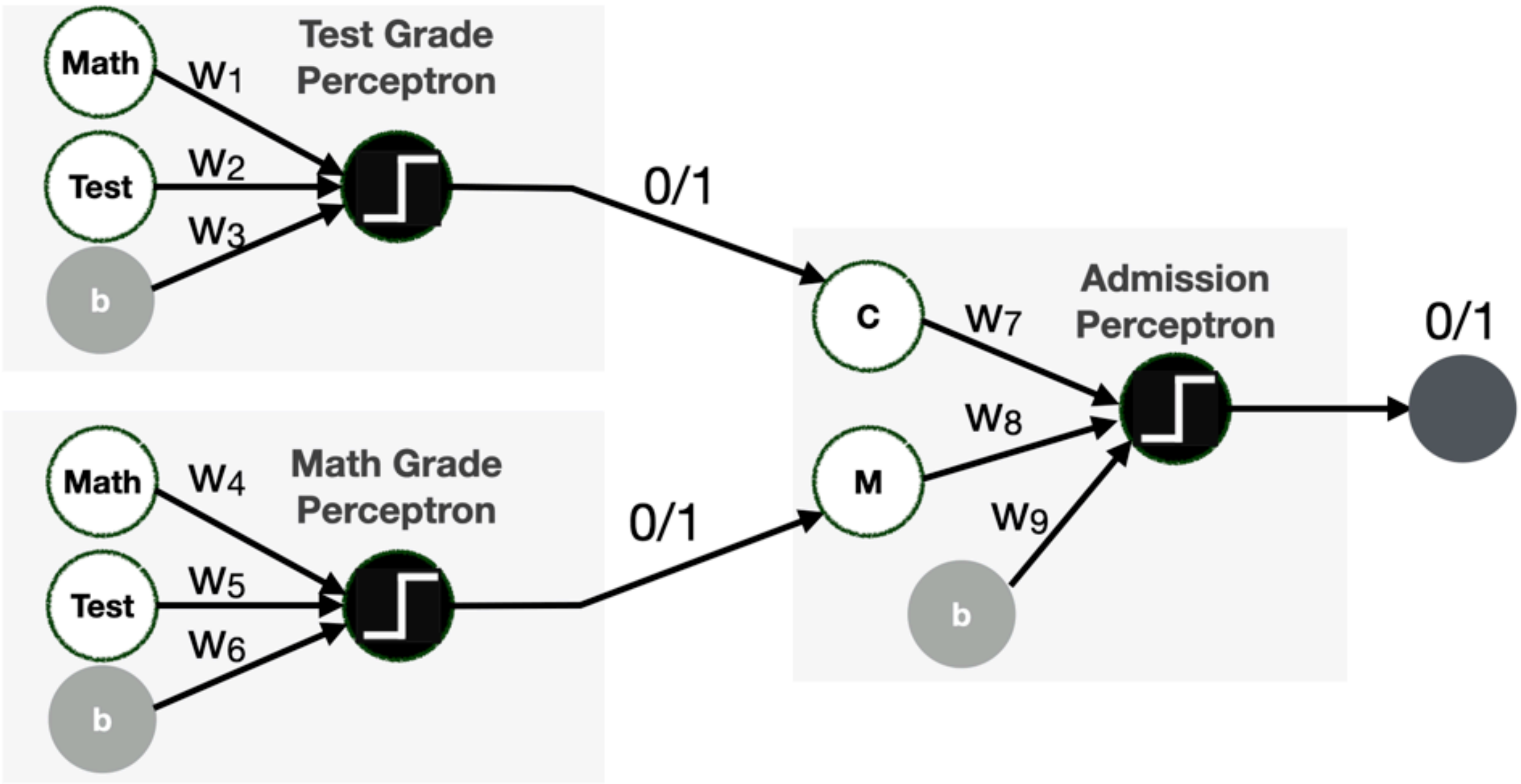
Remember our linear regression function?
It can be represented with a perceptron having a single input and a bias

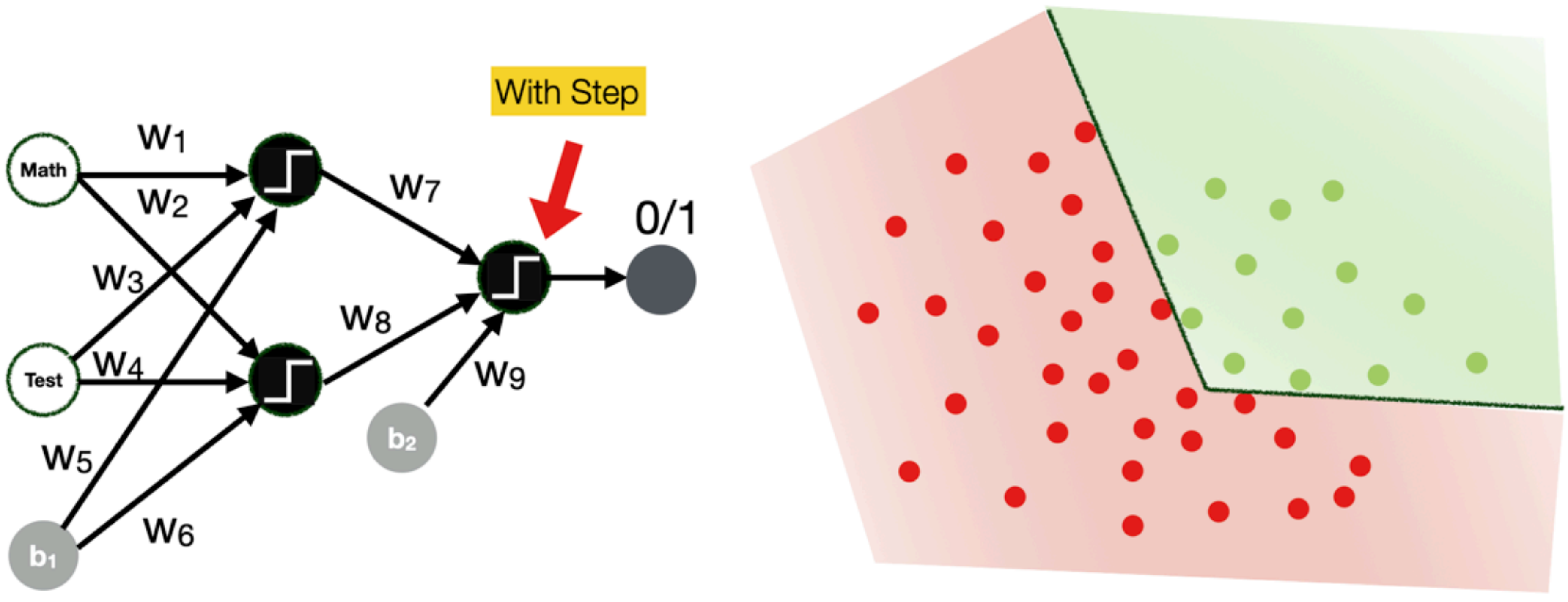


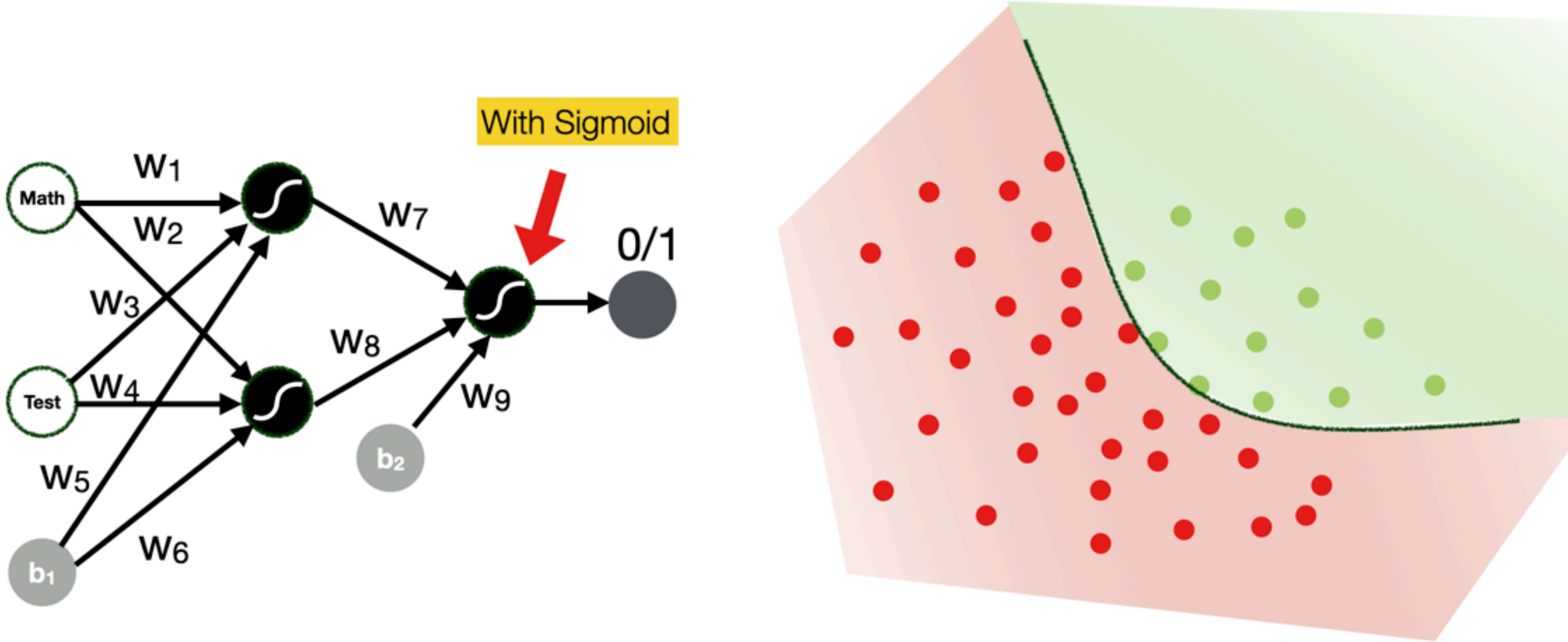

Cost = a + b Size
Remember our linear regression function?
It can be represented with a perceptron having
a single input and a single bias
Inputs
Weights
Activation function
Output
X1
X2
X3
1
3
-2
4
2
Bias
y = σ(3x1 − 2x2 + 4x3 + 2)
Step function
(discrete)
Output
1 if y is > 0
0 if y is <= 0
Input (weighted sum of values)
1
0
Sigmoid function
(continuous)
Output
σ(input) = 1 / (1 + e^−input)
Input (weighted sum of values)
1
0



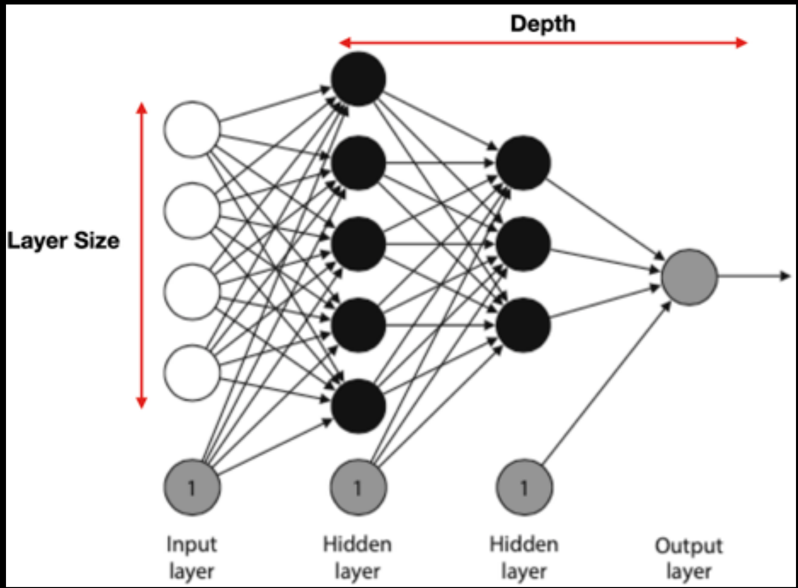

Test Grade Classifier
Math Grade Classifier
Admission Classifier


Math
Test
b
W1
W2
W3
Test Grade
Perceptron
Math
Test
b
W4
W5
W6
Math Grade
Perceptron
0/1
0/1
C
M
b
W7
W8
W9
Admission
Perceptron
0/1


With Step
W1
W2
W3
W4
W5
W6
W7
W8
W9
Math
Test
b1
b2
0/1

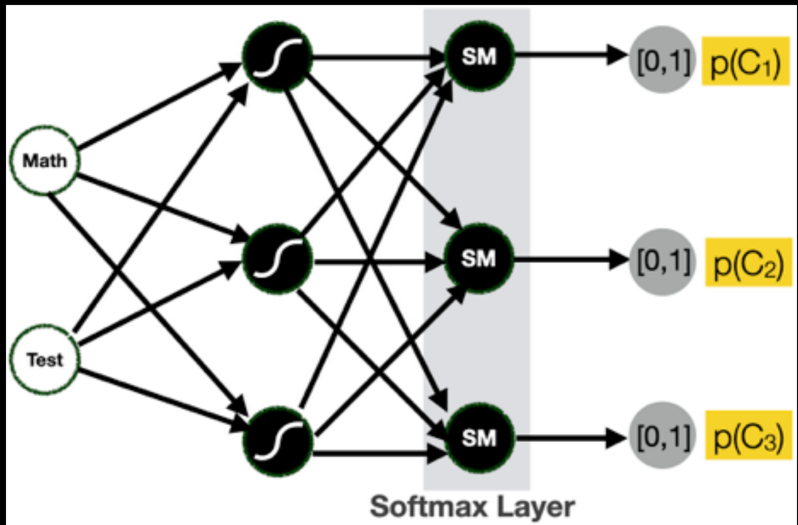

With Sigmoid
Math
Test
W1
W2
W3
W4
W5
W6
W7
W8
W9
b1
b2
0/1

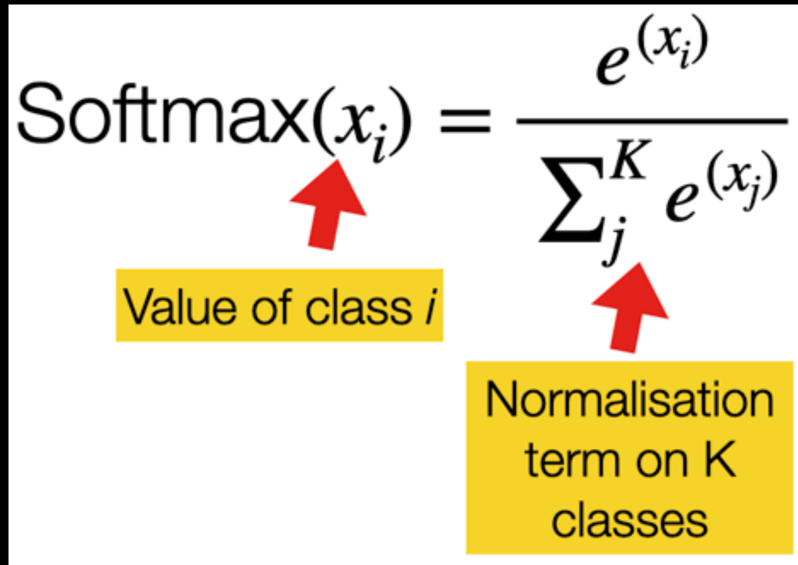
Fully connected Neural Network



- **Hyperparameters**
  - Learning rate
  - Number of epochs
  - Architecture
    - #layers, #nodes, activation functions
  - Batch vs. mini-batch vs. stochastic gradient descent
  - Regularization parameters:
    - Dropout probability p

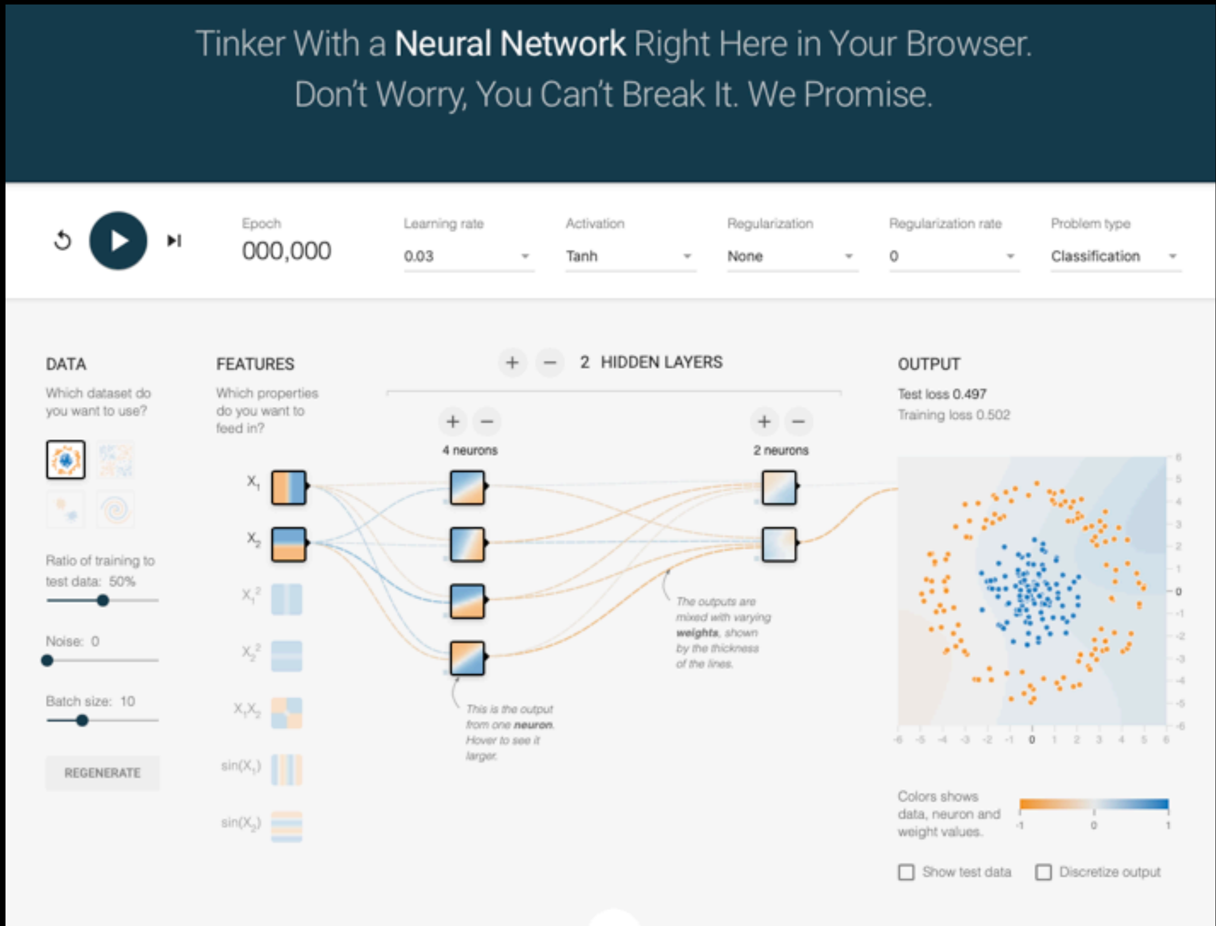
Classifying into multiple classes - *Softmax* function





- Return a probability for each class
  - example C1= ADMITTED, C2 = NOT ADMITTED, C3 = NEW TEST
  - p(C1) = 0.37, p(C2) = 0,21, p(C3) = 0,42
- We use the *Softmax* activation function for the output layer

Tensorflow Playground



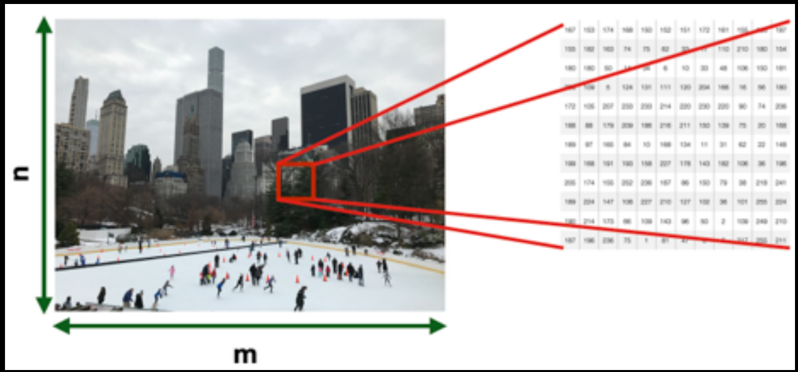
Machine Learning and Images

What do you see?

This is what a computer “sees”

167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187
155	182	163	74	75	62	33	17																																	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33																																	34	6	10	33	48	106	150	181
206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180
172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206
188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168
189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196
205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241
189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210
187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211
167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187
155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181
206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180	206	109	5	124	131	111	120	204	166	16	56	180
172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206	172	105	207	233	233	214	220	230	220	90	74	206
188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168	188	88	179	209	186	216	211	150	139	75	20	168
189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148	189	97	165	84	10	168	134	11	31	62	22	148
199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196	199	168	191	193	158	227	178	143	182	106	36	196
205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241	205	174	155	252	236	187	86	150	79	38	218	241
189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224	189	224	147	108	227	210	127	102	36	101	255	224
190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210	190	214	173	66	109	143	96	50	2	109	249	210
187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211	187	196	236	75	1	81	47	0	6	217	255	211
167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187	167	153	174	168	150	152	151	172	161	155	156	187
155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154	155	182	163	74	75	62	33	17	110	210	180	154
180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181	180	180	50	14	34	6	10	33	48	106	150	181

Images



- Each pixel in an image is a *feature*
  - numerical
    - 0 or 1 for *Black and White*
    - Between 0 and 255 for *greyscale*
    - 16M values for *RGB*
- Dimensionality → *n x m*

Computer Vision

Building algorithms that can “understand” the content of images and use it for other applications

- It is a “Strong AI” problem
  - signal-to-symbol conversion
  - The **semantic gap**

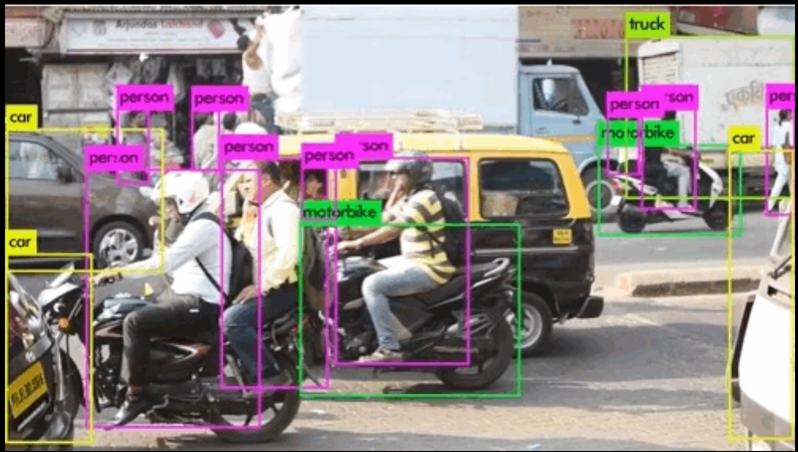
A general-purpose vision system **requires**

- Flexible, robust visual representation
- Updated and maintained
- Reasoning
- Interfacing with attention goals, and plans

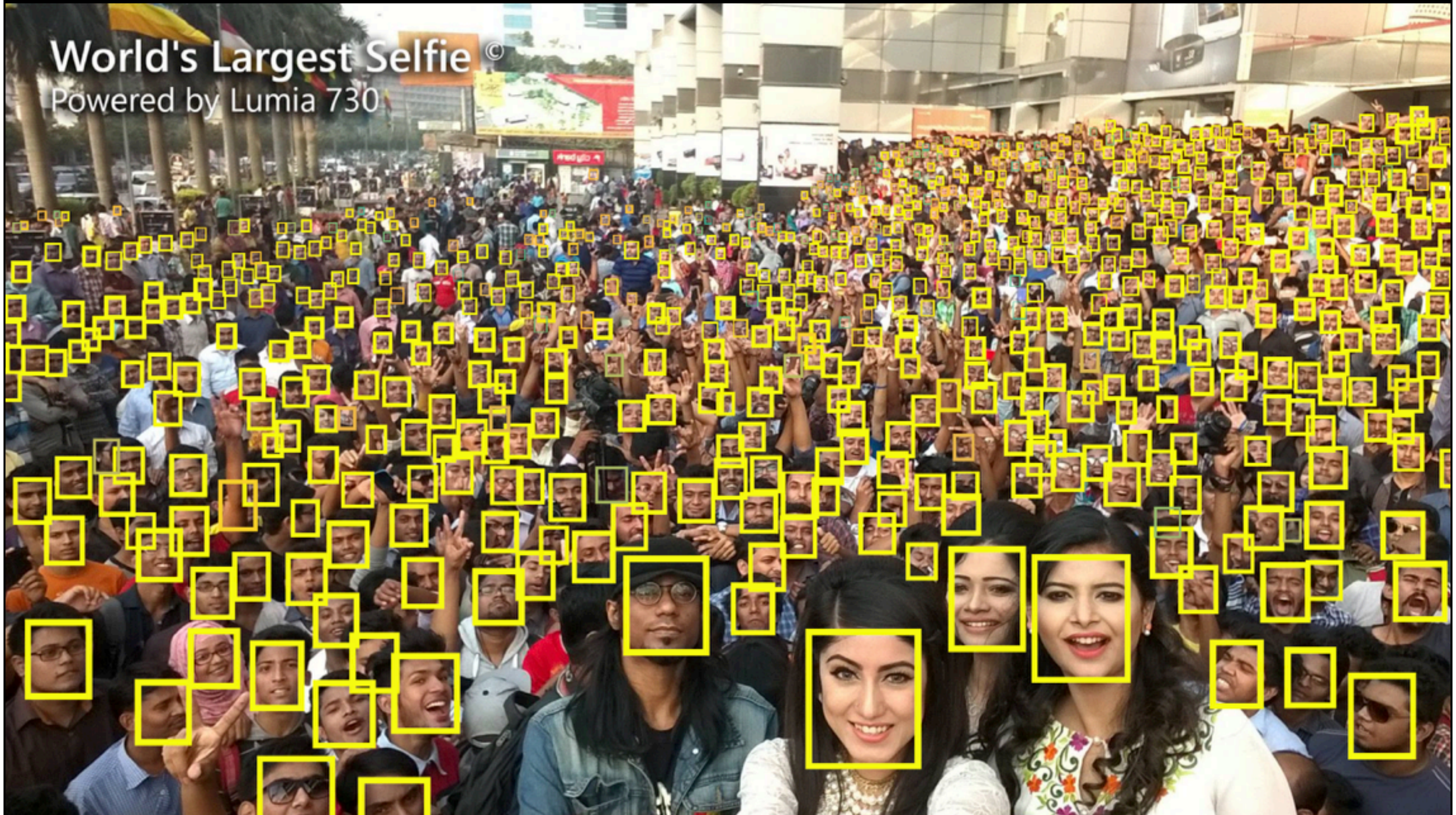
What specific tasks can we train a CV system to perform?


Is this a flag?
(Recognition / Classification)
TRUMP


Where are the people?
(Recognition/Classification)
TRUMP
TRUMP


truck
car
car
car
person
motorbike
motorbike


person
bicycle


World's Largest Selfie ©
Powered by Lumia 730


Is this Jeff?
(Identification)
TRUMP
TRUMP

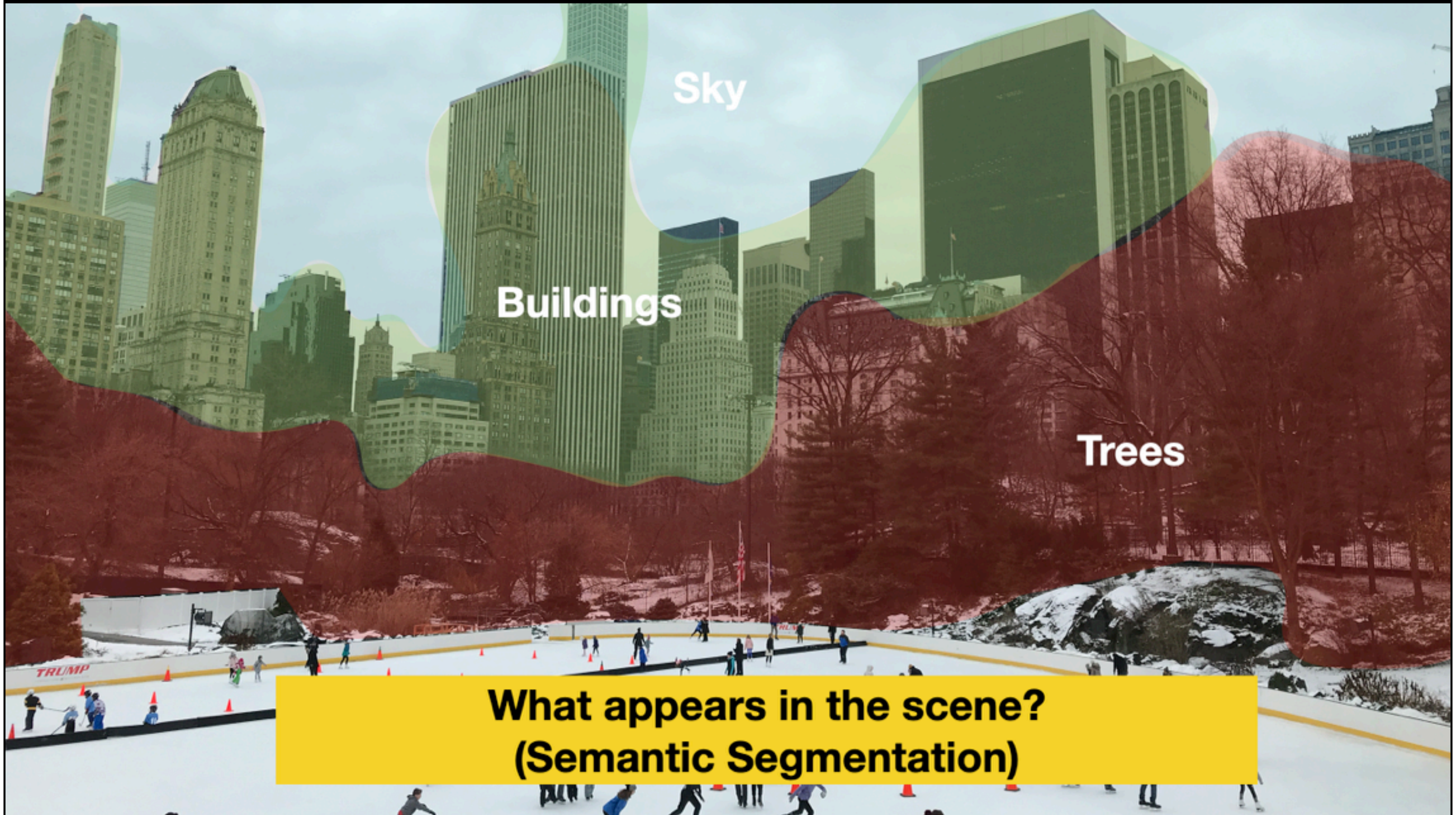

WA USA WASHINGTON
DRIVER LICENSE
FEDERAL LIMITS APPLY
4d LIC# WDLABCD456DG 9 CLASS
DONOR
1 TALBOT
2 LIAM R.
3 DOB 01/06/1958
4a ISS 01/06/2015
8 123 STREET ADDRESS
YOUR CITY WA 99999-1234
15 SEX M
16 HGT 5'-08"
12 RESTRICTIONS B
18 EYES BLU
17 WGT 165 lb
9a END L
4b EXP 08/12
5 DD WDLABCD456DG1234567XX1101
Veteran

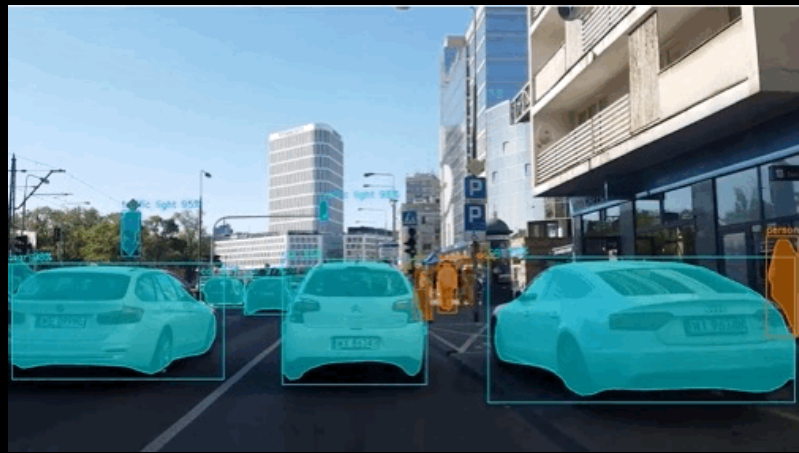

対象人物
対象人物
対象人物
対象人物


Is this the Wollman Rink?
(Identification)
TRUMP

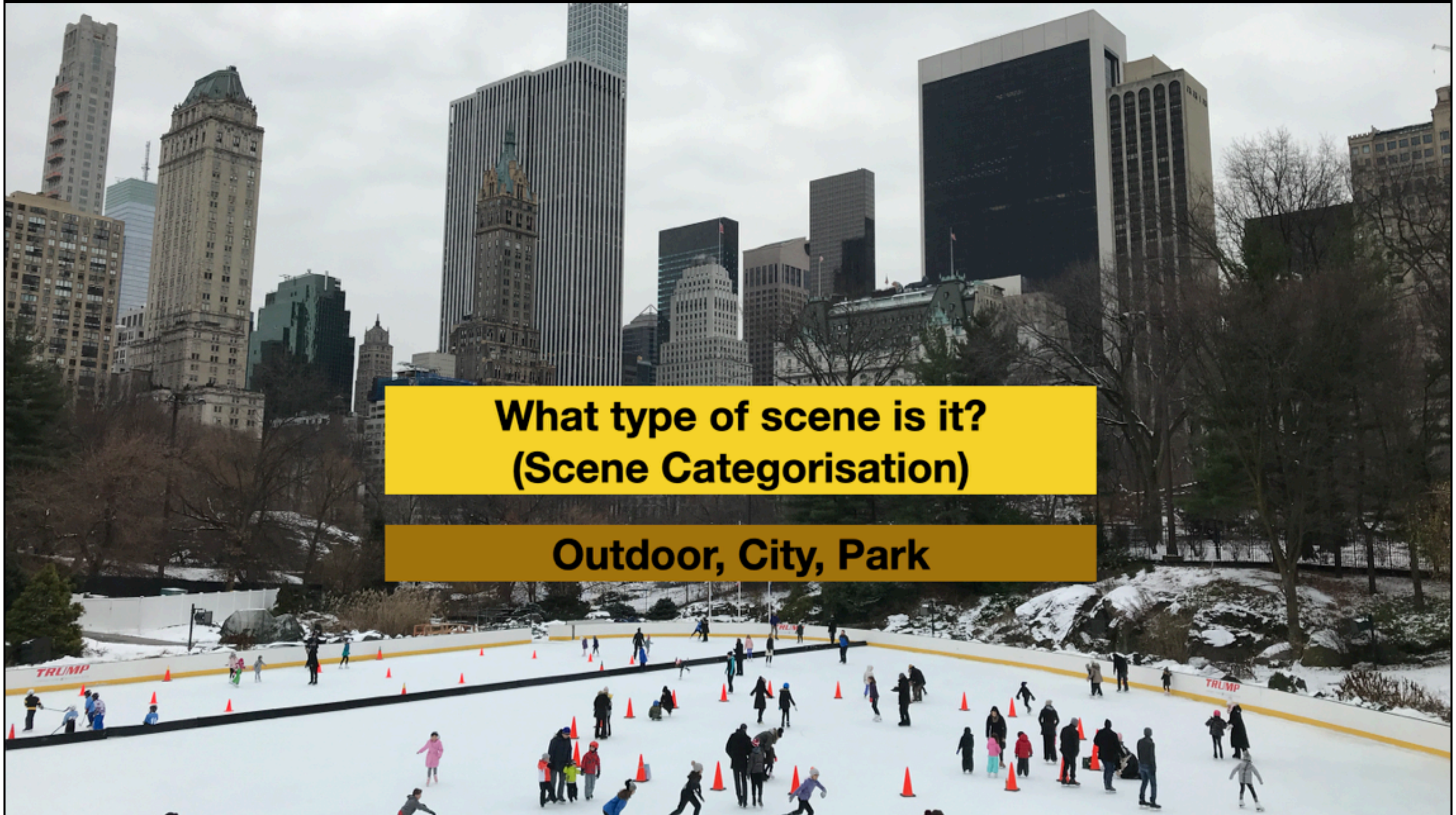

8:00
Google Lens
Coit Tower
Text selection
Coit Tower
4.4 2287 Reviews


Google Lens
DELFT
DESIGN
GUIDE
PERSPECTIVES - MODELS - APPROACHES
BISPUBLISHERS
Translate
Text
Search
Homework
Delft Design Guide: Design Strategies and Methods
Book
Search

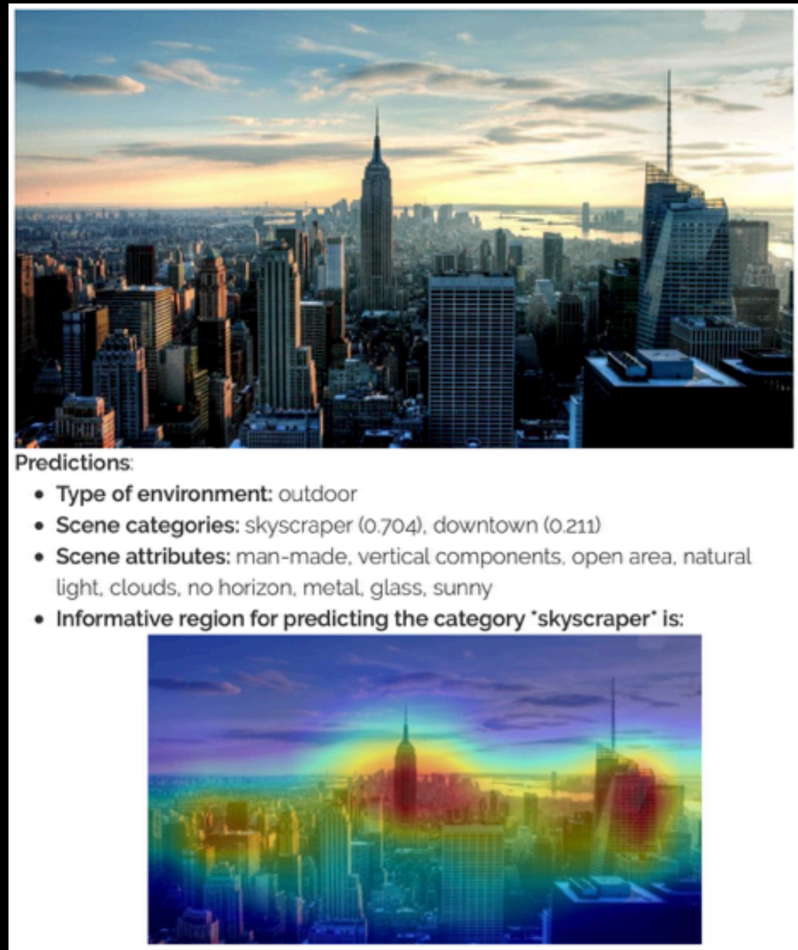

Sky
Buildings
Trees
TRUMP
What appears in the scene?
(Semantic Segmentation)



Project Sunroof


What type of scene is it?
(Scene Categorisation)
Outdoor, City, Park
TRUMP


darkest forest path
wintering forest path
greener forest path
rocky coast
misty coast
sunny coast


Predictions:
Type of environment: outdoor
Scene categories: skyscraper (0.704), downtown (0.211)
Scene attributes: man-made, vertical components, open area, natural light, clouds, no horizon, metal, glass, sunny
Informative region for predicting the category *skyscraper* is:

Hugging Face | Search models, datasets, users... | Models | Datasets | Spaces | Docs | Solutions | Pricing

Spaces: OFA-Sys / **OFA-Image_Caption** | like 12 | Running

App | Files and versions

OFA-Image_Caption

Gradio Demo for OFA-Image_Caption. Upload your own image or click any one of the examples, and click "Submit" and then wait for the generated caption.

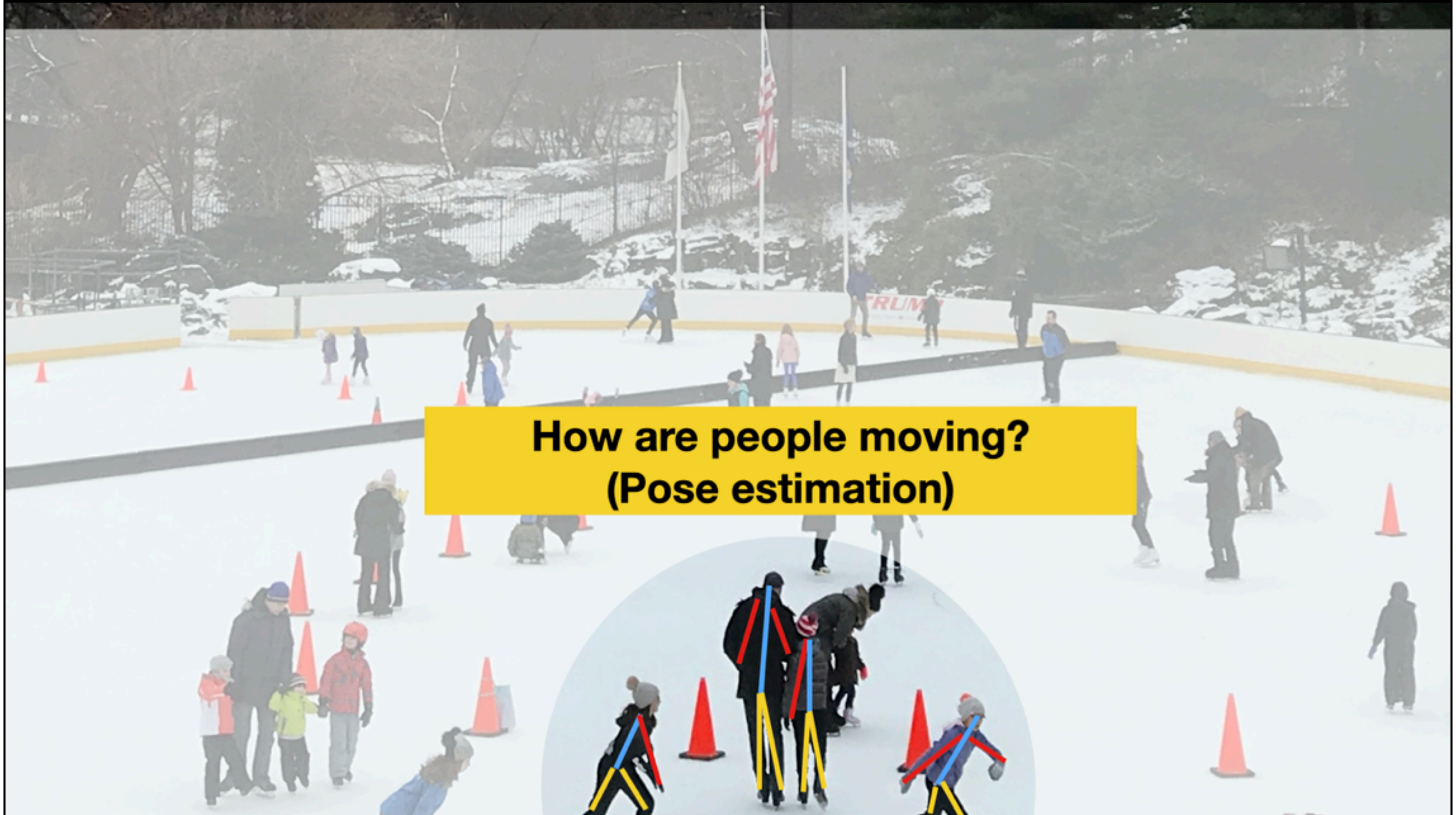
Image

Caption 8.28s

people skating on a rink in a city park with skyscrapers

Clear | Submit


What are these people doing?
(Activity / event recognition)
Skating
TRUMP
TRUMP

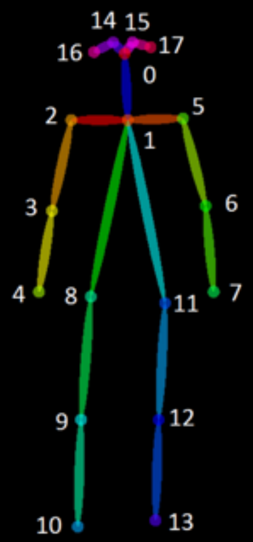

How are people moving?
(Pose estimation)

Stereolabs ZED Camera

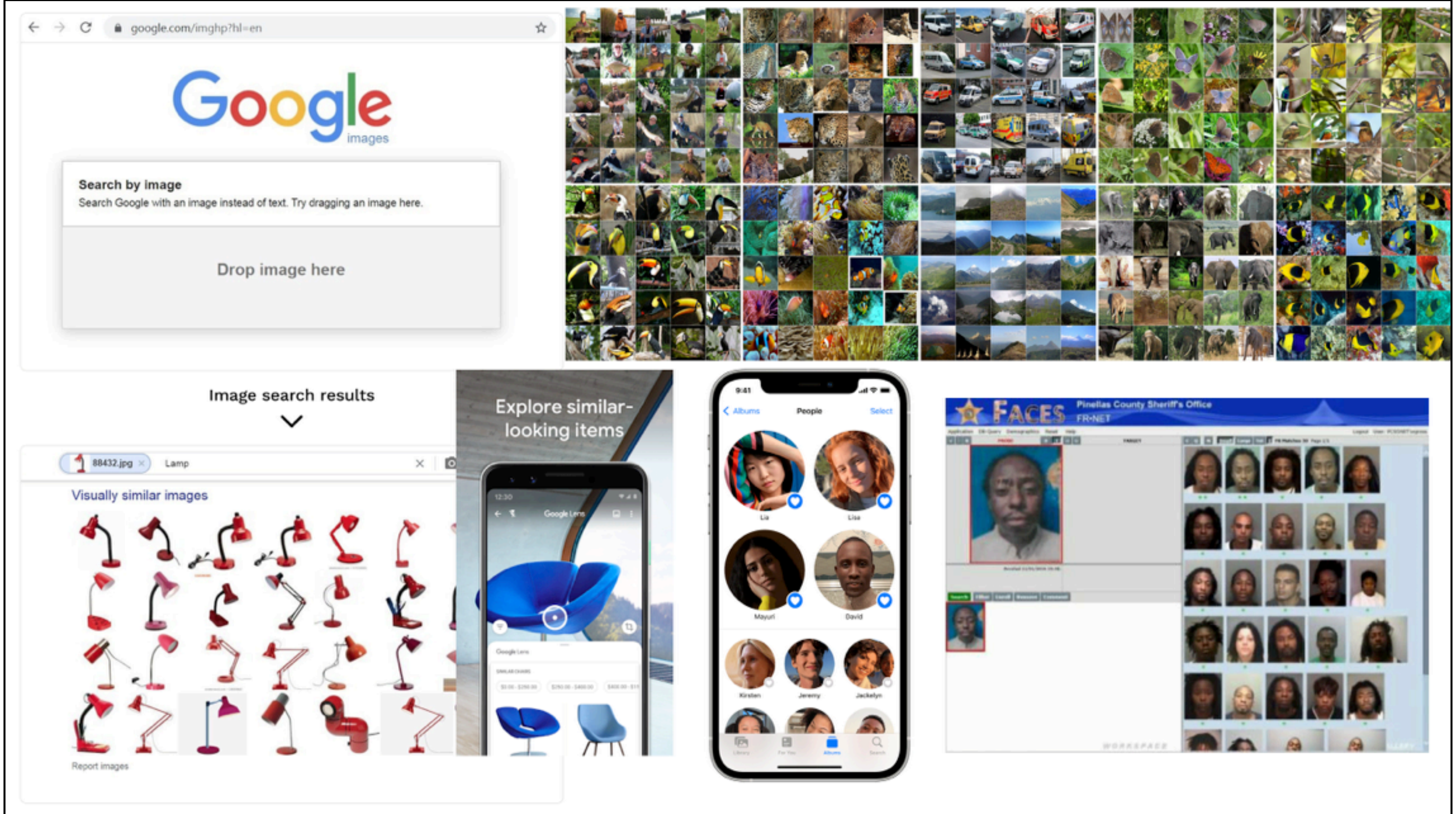
3D Object Detection

Body tracking

Positional tracking




Are these images of the same person?
(Image / Face Similarity)
Bonus if you guess the movie!


google.com/imghp?hl=en
Google
images
Search by image
Search Google with an image instead of text. Try dragging an image here.
Drop image here
Image search results
88432.jpg
Lamp
Visually similar images
Report images
Explore similar-looking items
Google Lens
9:41
Albums
People
Select
Lia
Lisa
Mayuri
David
Kirsten
Jeremy
Jackelyn
FACES
Pinellas County Sheriff's Office
FR-NET

Machine Learning for Design

Lecture 3

Machine Learning for Images. *Part 1*

Credits

CMU Computer Vision course - Matthew O'Toole.

Grokking Machine Learning. Luis G. Serrano. Manning, 2021